

**Tolloid metalloproteases implicated in dorsal-ventral
patterning and extra-cellular matrix activation in
*Xenopus laevis***

Timothy John Geach

A Thesis Submitted for the Degree of Doctor of Philosophy
University of London
2005

Department of Anatomy & Developmental Biology
University College London
London

UMI Number: U593584

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U593584

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

Tolloid (Tld) metalloproteases are zinc dependent extra-cellular endopeptidases that have numerous roles during embryonic development. All Tlds have a highly conserved N-terminal protease domain and an array of C terminal CUB and EGF-like domains thought to play a role in substrate interactions. In *Xenopus laevis* three members of the Tld family have been identified, BMP-1/Tld, Xolloid and Xolloid-related. All modulate dorsal-ventral patterning of the developing embryo by cleaving the dorsalising factor Chordin, preventing it from binding and inhibiting of the signalling molecule BMP-4. Biochemical studies of mammalian Tlds have identified a wide range of substrates, many involved in formation of the extra-cellular matrix. BMP-1/Tld is identical to pro-collagen C-proteinase, an enzyme that removes the C-terminal pro-peptide of procollagen types 1, 2 and 3, the N-terminal pro-peptide of procollagen type 11 and both the N- and C-terminal pro-peptides of procollagen type 5. It also activates lysyl-oxidase (lox), an enzyme that plays an essential role in collagen maturation. In addition, BMP-1 has been implicated in the proteolytic activation of biglycan, endorepellin, myostatin, osteoglycin and the α -3 and γ -2 chains of Laminin-5. In this thesis I identify and describe the expression for *Xenopus* homologues of *procollagen 3 α 1*, *5 α 1*, *5 α 2* and *11 α 1*, *biglycan* and a *laminin* α -chain like gene. I have also identified three members of the *lox* family and characterise their role during early *Xenopus* development. In addition, using a domain deletion approach, I determine the C-terminal CUB domains of Xolloid that are required for cleavage of Chordin. Finally, I study the potential role of endodermin, which displays significant homology to α -2 macroglobulin, as an inhibitor of Tld metalloproteases.

Table of Contents

Abstract	2
Table of Contents	3
List of Figures and Tables	8
Abbreviations	11
Acknowledgements	14
Chapter 1 General Introduction	
1.1 <i>Xenopus laevis</i> as a model organism	15
1.2 Early patterning of the <i>Xenopus</i> embryo	16
1.2.1 Establishment of the dorsal-ventral axis in <i>Xenopus</i> embryos	16
1.2.2 Mesoderm induction	17
1.2.3 The organiser and dorsalisation	20
1.2.4 Patterning the marginal zone - Bone morphogenetic proteins	22
1.3 The Tolloid Metalloproteases	24
1.4 The Tld protein arrangement	25
1.4.1 The Astacin protease domain	25
1.4.2 The pro-peptide domain and activation of protease activity	26
1.4.3 CUB and EGF-like domains	39
1.4.3.1 CUB domains	39
1.4.3.2 EGF-like domains	31
1.5 Cleavage recognition sites	32
1.6 The Tlds in embryonic development	33
1.7 The Tlds in Collagen formation	40
1.8 Other substrates of the Tlds	43
1.9 Summary	44
Chapter 2 Materials and Methods	
2.1 Materials and Solutions	46
2.1.1 DNA constructs and PCR primers	46
2.1.2 Solutions and media	47
2.2 Methods	51
2.2.1 <i>Xenopus</i> techniques	51
2.2.2 Embryo manipulations	52

2.2.2.1	Animal cap removal	52
2.2.2.2	DMZ and VMZ removal	52
2.2.2.3	Oocyte collection	53
2.2.2.4	Microinjection of mRNA	53
2.2.3	DNA/RNA techniques	53
2.2.3.1	Agarose gels	54
2.2.3.2	Restriction digests	54
2.2.3.3	Cleaning of digests	54
2.2.3.4	Synthetic mRNA transcription	54
2.2.2.5	RT-PCR	55
2.2.2.6	PCR for deletion constructs	57
2.2.2.7	Sequencing	58
2.2.4	Protein Analysis	58
2.2.4.1	Cell Culture and transfection	58
2.2.4.2	<i>In vitro</i> translation	59
2.2.4.3	Western Blot	59
2.2.5	Wholemount <i>in situ</i> hybridisation	60
2.2.5.1	Probe synthesis	60
2.2.5.2	<i>In situ</i> hybridisation	60
2.2.6	Wholemount antibody staining	63
2.2.7	Histology	63
2.2.7.1	Wax embedding of embryos	63
2.2.7.2	Sectioning	64
2.2.7.3	Counter staining	64
2.2.8	Electron microscopy	65
2.2.9	Photography	65
Chapter 3	Identification and characterisation of three members of the Lysyl oxidase family expressed in the notochord of developing <i>Xenopus laevis</i> embryos	
3.1	Introduction	66
3.1.1	Lysyl oxidase	66
3.1.2	Lox-like enzymes	67
3.1.3	Inhibitors of Lox function	71

3.1.4	Summary	71
3.2	Results	73
3.2.1	Identification of three <i>lox</i> genes during <i>X. laevis</i> development	73
3.2.2	<i>Xlox</i> , <i>Xloxl-1</i> and <i>Xloxl-3</i> are expressed in the notochord during <i>X. laevis</i> embryogenesis	78
3.2.3	Inhibition of LOX function by β -APN treatment results in defects consistent with collagen malformation and does not affect A-P axis formation.	84
3.2.4	<i>X. laevis</i> embryos are sensitive to β -APN throughout development	94
3.2.5	The notochordal sheath and somitic muscle is severely disrupted in β -APN treated embryos	100
3.2.6	Injecting <i>Xlox</i> and <i>Xloxl-1</i> mRNA has no effect on embryonic patterning	104
3.3	Discussion	109
3.3.1	<i>X. laevis</i> <i>lox</i> genes	109
3.3.2	Function of <i>lox</i> genes in early <i>X. laevis</i> development	110
3.3.3	<i>X. laevis</i> Lox function in notochord and somite development	111
3.3.4	Inhibition of Lox function using β -APN	112
3.3.5	Summary	114
Chapter 4 The first and second CUB domains are required for Xld to process Chd		
4.1	Introduction	115
4.1.1	Tld metalloprotease activity	115
4.1.2	CUB domains	116
4.1.3	EGF-like domains	117
4.1.4	Summary	118
4.2	Results	119
4.2.1	Both CUB1 and CUB2 are required for the ventralising activity of Xld	119
4.2.2	Removal of both CUB4 and CUB5 does not enhance the activity of Xld	125
4.2.3	Both CUB1 and CUB2 of Xld are required to counteract co-injected Chd	128
4.2.4	The presence of CUB2 is required for secretion of Xld	134
4.2.5	Dominant negative Xld constructs	137
4.3	Discussion	144
4.3.1	CUB and EGF-like domains in Xld	144

4.3.2	Dominant-negative constructs	147
4.3.3	Summary	148
Chapter 5 Endodermin is not an inhibitor of Tld proteases		
5.1	Introduction	151
5.1.1	The Tlds during <i>Xenopus</i> gastrulation	151
5.1.2	Endodermin and α -2 Macroglobulin	152
5.1.3	Proteases inhibited by α -2M	153
5.1.4	Aims	156
5.2	Results	157
5.2.1	Expression of <i>edd</i> by whole mount <i>in situ</i> hybridisation	157
5.2.2	Edd does not counteract the effect of <i>xbmp-1</i> over-expression, or its action on Chd.	160
5.2.3	Edd does not bind XBMP-1	166
5.2.4	α -2 macroglobulin does not inhibit cleavage of Chd in Xlr conditioned media	167
5.3	Discussion	170
5.3.1	Edd in <i>Xenopus</i> development	170
5.3.2	Potential roles of Edd and α -2M as protease inhibitors during embryogenesis	172
5.3.3	Other potential roles for Edd	174
5.3.4	Summary	174
Chapter 6 Identification of potential substrates for the <i>Xenopus</i> Tlds		
6.1	Introduction	176
6.1.1	Substrates of the Tlds	176
6.1.2	Aims	178
6.2	Results	180
6.2.1	Fibrillar procollagens expressed in early <i>Xenopus</i> development	180
6.2.2	<i>Xbgn</i> is expressed in the notochord and somites after gastrulation	183
6.2.3	A Laminin α -like chain is expressed in the notochord	189
6.3	Discussion	196
6.3.1	Fibrillar Collagens	196
6.3.2	XBgn	197
6.3.3	Laminin α -like	199

6.3.4	Summary	201
Chapter 7 General Discussion		
7.1	The role of Tld metalloproteases	202
7.2	The role of C-terminal CUB domains in Xld	203
7.3	Do inhibitors regulate Tld activity?	206
7.4	Putative Tld substrates	208
7.4	Closing thoughts	214
References		217
Appendices		236

List of Figures

Chapter 1	General Introduction	
Figure 1.1	Domains of the Tld metalloproteases	28
Figure 1.2	The “Met-turn”	28
Table 1.1	Cleavage Sites of the Tlds	34
Figure 1.3	Tld cleaves Chd/Sog releasing bound BMP-4/Dpp	38
Figure 1.4	Assembly of mature collagen fibrils in the ECM	38
 Chapter 2	 Materials and Methods	
Table 2.1	ESTs and PCR primers used in this thesis	46-47
 Chapter 3	 Identification and characterisation of three members of the Lysyl oxidase family expressed in the notochord of developing <i>Xenopus laevis</i> embryos	
Figure 3.1	A schematic representation of the Lox family of proteins	70
Figure 3.2	Deduced protein sequences and alignments for the <i>Xenopus</i> Lox family	75
Figure 3.3	Temporal expression of <i>xlox</i> , <i>xloxl-1</i> , and <i>xloxl-3</i> as determined by RT-PCR	80
Figure 3.4	Expression of <i>Xlox</i> , <i>Xloxl-1</i> , and <i>Xloxl-3</i> determined by wholemount <i>in-situ</i> hybridisation	83
Figure 3.5	Treatment of <i>Xenopus</i> embryos with 0.25mM β -APN	86
Figure 3.6	Sections through the length of stage 43/44 embryos treated with 0.25mM β -APN	88
Figure 3.7	Immunolocalisation of muscle and notochord markers in embryos treated with β -APN	91
Figure 3.8	Muscle is present in anterior regions of β -APN treated embryos at stage 28 but absent at stage 41	93
Figure 3.9	β -APN treatment during different stages of development	96
Figure 3.10	Notochord defects appear more posteriorly the later in development β -APN is applied	99
Figure 3.11	T.E.M analysis of collagen fibres in the notochordal sheath and	103

	somatic muscle in β -APN treated embryos	
Figure 3.12	β -APN does not prevent the elongation of isolated DMZ's, nor the formation of muscle and notochord in these explants	106
Figure 3.13	β -APN does not inhibit elongation of activin treated animal caps	106
Figure 3.14	Over-expression of <i>Xlox</i> and <i>Xloxl-1</i>	108
Chapter 4	The second CUB domain of Xld is required by the protease to process Chd	
Figure 4.1	Schematic representation of CUB deletion constructs	121
Figure 4.2	Ectopic expression of each deletion construct in whole embryos	124
Figure 4.3	<i>In vivo</i> translation of Xld deletion constructs	124
Figure 4.4	<i>In vitro</i> translation of <i>XldC1-2</i> and <i>XldC1-e</i>	124
Figure 4.5	Ventral injection of <i>Xld</i> and <i>XldC1-3</i>	127
Figure 4.6	Optimisation of <i>Xld</i> and <i>XldC1-3</i> injections	127
Figure 4.7	Dorsal and Ventral injection of <i>Xld</i> and <i>XldC1-3</i>	130
Figure 4.8	Coinjection of <i>Xld</i> deletion constructs and <i>chd</i> mRNA	133
Figure 4.9	Secretion of Xld deletions	136
Table 4.1	Predicted molecular weights of Xld CUB deletions	136
Figure 4.10	Schematic diagram of putative dnXld CUB deletions	139
Figure 4.11	Injection of mRNA for dnCUB deletions	143
Figure 4.12	Summary of results obtained with each CUB deletion construct	150
Chapter 5	Endodermin is not an inhibitor of Tld proteases	
Figure 5.1	A comparison of the bait region of Edd and α -2M shows little conservation	155
Figure 5.2	Spatial expression of <i>edd</i> during embryogenesis	159
Figure 5.3	Comparison of the expression pattern of <i>edd</i> and <i>chd</i> in gastrulae	159
Figure 5.4	Edd does not inhibit the ventralising activity of BMP-1	162
Figure 5.5	Edd does not prevent BMP-1 from blocking the secondary axis inducing activity of Chd	165
Figure 5.6	Edd does not prevent BMP-1 from blocking the dorsalising activity of Chd in VMZs	165
Figure 5.7	Edd does not bind to BMP-1	169

Figure 5.8	α -2M does not inhibit Xlr	169
Chapter 6	Identification of potential substrates for the <i>Xenopus</i> Tlds	
Figure 6.1	Expression of fibrillar procollagens in early <i>Xenopus</i> development	182
Figure 6.2	Amino acid sequence of XBgn and XBgn-EST	185
Figure 6.3	Expression of <i>Xbgn-EST</i>	188
Figure 6.4	Alignment of predicted protein sequence of <i>X. tropicalis</i> and <i>X. laevis</i> laminin α -like	191
Figure 6.5	Expression of <i>laminin α-like</i>	195
Chapter 7	General Discussion	
Figure 7.1	Summary of Tld function	217

Abbreviations

1,10 PT	1, 10 Phenanthroline
α -2M	α -2 Macroglobulin
ADAM	Adamalysins A Disintegrin and Metalloproteinase
ADAMTS	Adamalysin with thrombosporin domain
β -APN	β -Aminopropionitrile
A-P	Anterior-posterior
Bgn	Biglycan
BMP	Bone morphogenetic protein
Chd	Chordin
CUB	Complement serum proteins Clr/Cl _s , Uegf, BMP-1
CR	Cysteine-rich domain (e.g. of Chd)
DAI	Dorso-Anterior Index
DLZ	Dorso-lateral Zone
DMZ	Dorsal Marginal Zone
Dpp	Decapentaplegic
ECM	Extracellular Matrix
Edd	Endodermin
EGF-like	Epidermal growth factor-like domain
EST	Expressed sequence tag
EM	Electron microscopy
FGF	Fibroblast growth factor
HCG	Human chorionic gonadatrophin
LG-domain	Laminin globular domain
Lox	Lysyl oxidase
LOX	Lysyl oxidase catalytic domain
Loxl	Lysyl oxidase-like
LTQ	Lysyl-tyrosylquinone
MASP	Mannose binding protein serine associated protease
MBP	Mannose binding protein
MBT	Midblastula transition
MMP	Matrix Metalloprotease

mTld	Mammalian Tolloid
mTll	Mammalian Tolloid-like
mRNA	Messenger Ribonucleic acid
NAM	Normal amphibian media
NC	Non-collagenous domain (in Procollagens)
ODC	Ornithine decarboxylase
Ogn	Osteoglycin
PC	Proprotein convertase
PCP	Procollagen C-proteinase
PCPE	Procollagen C-proteinase enhancer
PNP	Procollagen N-proteinase
ProCol α	Procollagen alpha chain
RT-PCR	Reverse transcriptase – polymerase chain reaction
Scr	Screw
Sog	Short gastrulation
SRCR	Scavenger Receptor Cysteine Rich domain
T.E.M	Transmission Electron microscopy
TGF- β	Transforming growth factor- β
Tld	Tolloid
Tlk	Tolkin
Tll	Tolloid-like
UTR	Untranslated region
UV	Ultraviolet
Tsg	Twisted gastrulation
VLZ	Vento-lateral Zone
VMZ	Ventral Marginal Zone
XBgn	<i>Xenopus</i> Biglycan
XBMP	<i>Xenopus</i> bone morphogenetic protein
Xld	Xolloid
Xlr	Xolloid-related
Xnr	<i>Xenopus</i> Nodal-related
Zer	Zerknüllt
zTld	Zebrafish Tld

Units

bp	base pair
°C	degrees Celcius
g	gram
kb	kilo base
kDa	kilo Dalton
l	litre
min	minute
M	Molar concentration
sec	seconds

A note on conventions used in the text.

When referring to a gene/mRNA its name or abbreviation appears in *italics*, as opposed to normal text for the product/protein. In some cases amino acids are abbreviated to their standard three, or single, letter code.

Acknowledgements

Firstly, I am indebted to my supervisor Dr. Les Dale, for his time, support, enthusiasm and patience, without which the completion of this thesis would have been impossible. Thanks must also go to Professor Anne Warner for her help and always having an open door, particularly during difficult times, and to Christelle Devader, for making work in the Lab so much fun.

Thank you to those people who provided me with much needed technical help in order to complete aspects of this work, especially Mary Rahman and Dr. Rachael Webb, to my family for helping me get here in the first place and to anyone who listened my rants about Tolloid metalloproteases and *Xenopus* over a beer.

Finally, I would like to give a special thank you to my girlfriend Louise for putting up with me during this last year of work and providing me with coffee in bed every morning. Yes, we can finally go on holiday now!

I gratefully acknowledge the financial support of the BBSRC and the physiological support of caffeine and EC 1.1.1.1.

This Thesis is dedicated to my teachers

Chapter 1

General Introduction

1.1 *Xenopus laevis* as a model organism

The South African clawed frog *Xenopus laevis* is one of only a few “model” organisms used for the vast majority of research into embryonic development. Historically, embryologists had used many different species of amphibia for their studies, essentially because they were available in local ponds and waterways. One limitation of this approach is that experiments were restricted to the breeding season and this would probably have confined amphibians to the scientific backwaters but for the need to develop an effective pregnancy test in the first half of the 20th century (Gurdon and Hopwood, 2000). The discovery that *X. laevis* could easily be maintained in the laboratory and that females respond to human pregnancy hormone by ovulating, led to the adoption of this frog as the bioassay of choice in the early diagnosis of pregnancy (Elkan, 1938). It was also a boon for developmental biologists who were now able to induce egg laying, in large quantities, at all times of the year. As a consequence, *X. laevis* soon became the amphibian of choice for embryological experiments (Gurdon and Hopwood, 2000). The relatively large egg size (~1.5 mm) makes it ideal for traditional microsurgical experiments, while the ability to produce hundreds of synchronously developing embryos makes it beneficial to both molecular and biochemical studies. The main drawback of *X. laevis* being that it is not suited for genetic experiments, due to its long life cycle and pseudotetraploid chromosomal make up. Recently, the smaller but more genetically malleable *X. tropicalis* has been employed for genetic studies. This truly diploid organism allows transgenic animals to be made, using techniques first developed for

X. laevis, and both forward and reverse genetic screens have recently been initiated (Noramly et al., 2005).

1.2 Early patterning of the *Xenopus* embryo

1.2.1 Establishment of the dorsal-ventral axis in *Xenopus* embryos

The *Xenopus* oocyte develops an animal and vegetal polarity during oocyte maturation. The animal pole is more darkly pigmented and beneath the cortex resides a large nucleus, the germinal vesicle, while the vegetal pole contains a yolky cytoplasm and a pool of specific maternal mRNAs. Upon laying, the dorsal and ventral axes have yet to be specified and are only done so after fertilisation, when a process known as cortical rotation occurs (Gerhart et al., 1989). Cortical rotation happens before the first cellular cleavage and realigns the cortex of the egg in relation to the deeper cytoplasm, so that the sperm entry point is roughly opposite the future dorsal side. During rotation an array of microtubules form that allow the movement of molecules from the vegetal pole to the future dorsal side of the embryo. Blocking the formation of this microtubule network by UV irradiating the vegetal pole, prevents these molecules from being positioned in their correct regions. In these embryos there is a massive expansion of ventral cell fates and dorsal structures such as the notochord and somites are lost.

The molecules that translocate to the dorsal side during cortical rotation are known as “dorsal determinants”, as their absence precludes the development of dorsal structures (Moon and Kimelman, 1998). There is compelling evidence that suggests these determinants are components of the Wnt signalling pathway. Embryos depleted of maternal β -catenin, a protein that combines with the transcription factor Tcf-3 and

regulates gene expression in the nucleus upon stimulation of the Wnt pathway, resemble embryos that have been irradiated with UV light (Heasman et al., 1994). More recent observations have shown that the proteins dishevelled (Dsh) and glycogen synthase kinase binding protein (GBP), which function upstream of β -catenin, travel along the microtubule pathway to the dorsal side during cortical rotation (Miller et al., 1999; Weaver et al., 2003). Both Dsh and GBP are inhibitors of glycogen synthase kinase 3- β (GSK-3 β), an enzyme that phosphorylates β -catenin and targets it for degradation (Klein and Melton, 1996; Stambolic et al., 1996). It is therefore apparent that inhibition of GSK-3 β is necessary for development of the dorsal structures. A number of observations support this, firstly inhibition of GSK-3 β by lithium chloride (LiCl) results in embryos with overtly dorsalised phenotypes (Kao and Elinson, 1989). Secondly, constitutively activated GSK-3 β , not inhibited by Dsh, has a ventralising effect when injected into dorsal blastomeres and a dominant negative version gives rise to complete secondary axes in when expressed in the future ventral marginal zone (VMZ) (He et al., 1995). It is now apparent that we have a situation in the dorsal half of the embryo whereby inhibition of GSK-3 β stabilises β -catenin, which then moves to the nucleus and together with Tcf-3 regulates expression of genes required for dorsal development in the embryo.

1.2.2 Mesoderm induction

Once the dorsal-ventral axis has been established the embryo begins the process of forming the three different germ layers, ectoderm, mesoderm and endoderm. Mesoderm is formed around the equator of the embryo in a region known as the “marginal zone” in response to signals released by cells in the vegetal hemisphere. Work in the late 1960’s by Peter Nieuwkoop, using axolotl blastulae, demonstrated

that vegetal explants could divert the fate of animal cap blastomeres from ectoderm to mesoderm (Nieuwkoop, 1969). When in the 1980's this same principle was investigated in *Xenopus* embryos, it was apparent that cells from a small region in the dorsal side of the vegetal pole could induce mesoderm of dorsal character (notochord and somitic muscle), while the remaining cells induced ventral mesoderm (mesothelium and blood) (Dale and Slack, 1987; Dale et al., 1985). It was now apparent that there were two signalling centres in the vegetal pole, one inducing ventral and the other inducing dorsal mesodermal fates. The dorsal signalling centre was subsequently named the Nieuwkoop centre, in honour of Peter Nieuwkoop's initial discovery (Gerhart et al., 1989).

The molecular nature of the mesoderm-inducing signal is still open to some debate and a number of molecules, including members of the fibroblast growth factor (FGF) and transforming growth factor β (TGF- β) families, are able to mimic the effect of vegetal pole cells and induce mesoderm in animal cap assays. To fit the criteria for an endogenous mesoderm-inducing signal, a molecule must be expressed in the vegetal half of the embryo during blastulae stages, be able to mimic the mesoderm induction effect of vegetal cells in animal cap assays, and its inhibition must prevent the formation of mesoderm in the embryo.

FGFs were the first mesoderm inducing factors to be identified in *Xenopus* embryos (Kimelman and Kirschner, 1987; Slack et al., 1987). They usually induce mesoderm of ventral character in animal cap assays, although at high concentrations they may induce muscle (Kimelman and Maas, 1992). Notochord, the most dorsal mesodermal cell type, is never formed. However, it is unlikely that an FGF is responsible for the

mesoderm-inducing signal. First, FGFs are expressed in the mesoderm after the blastula stage, rather than being localised to the vegetal pole as expected of a mesoderm-inducing factor (Isaacs, 1997). Second, inhibition of FGF signalling, by injection of dominant-negative receptors, only disrupts posterior mesoderm formation, while extreme anterior mesoderm appears relatively unaffected (Amaya et al., 1991; Isaacs, 1997). FGFs may have a role in mesoderm formation, since some members of this family, such as eFGF, are expressed in the mesoderm during gastrula stages (Isaacs et al., 1995). They may be responsible for maintaining the mesoderm following its initial induction (Isaacs et al., 1994; Schulte-Merker and Smith, 1995).

The first member of the TGF- β family to be put forward as the primary mesoderm-inducing signal was activin, which can induce all types of mesoderm, with the exception of blood, when applied to animal caps (Asashima et al., 1990; Smith et al., 1990). High concentrations of activin induce dorsal-type mesoderm while low concentrations induce ventral-type mesoderm (Green and Smith, 1990). This was followed by a host of other TGF- β family members including Vg1, *derriere*, and members of the nodal-related (Xnr) family. All of these molecules are all localised to the vegetal hemisphere of *Xenopus* blastulae and induce mesoderm in a concentration dependent manner in animal caps (Dale et al., 1993; Jones et al., 1995; Joseph and Melton, 1997; Osada and Wright, 1999; Smith et al., 1995; Sun et al., 1999; Takahashi et al., 2000; Thomsen and Melton, 1993). Of these, perhaps the most likely molecules to be the mesoderm-inducing factor(s) of the vegetal pole are member(s) of the Xnr family. Nodal was first identified in mice and embryos that carry a mutation in this gene form little mesoderm and die during gastrulation (Conlon et al., 1994). In *Xenopus*, there are presently six nodal-related genes that

have been identified, five of which induce mesoderm in animal caps and are expressed in the vegetal hemisphere of blastula stage embryos (Agius et al., 2000; Jones et al., 1995; Takahashi et al., 2000). While inhibition of individual Xnrs has yet to be done, the gross inhibition of all can be achieved using the C-terminal fragment of the BMP, Wnt and nodal inhibitor, cerberus (Cer) (Agius et al., 2000; Piccolo et al., 1999). The C-terminal specifically inhibits Xnrs in *Xenopus* and blocks mesoderm formation when injected into *Xenopus* embryos. Thus Xnrs meet all three criteria required for the endogenous mesoderm-inducing factor.

It is also apparent that a maternal T-box transcription factor, VegT, plays a central role in initiating expression of genes encoding the mesoderm-inducing signal (Hill, 2001). VegT is located to the vegetal hemisphere and depletion of transcripts using antisense oligonucleotides prevents formation of mesoderm and vegetal pole explants from these embryos do not induce mesoderm in animal cap conjugates (Zhang et al., 1998). Moreover, VegT regulates the expression of TGF- β family members implicated in mesoderm induction, including *Xnr1*, *Xnr2*, *Xnr4* and *derriere* (Kofron et al., 1999). While VegT activates the *Xnr* genes at low levels throughout the vegetal hemisphere, in the Nieuwkoop centre it acts in combination with β -catenin, inducing higher levels of expression and a gradient of dorsal to ventral Xnr activity (Agius et al., 2000; Hyde and Old, 2000). This provides further support for the proposed role of Xnr family as the endogenous mesoderm-inducing factors.

1.2.3 The organiser and dorsalisation

The mesoderm-inducing signal released by the Nieuwkoop centre induces the dorsal marginal zone (DMZ) to develop into extreme dorsal cell types such as the notochord

(Dale and Slack, 1987). In *Xenopus* embryos the DMZ is first apparent with the formation of the dorsal blastopore lip an arc of pigmentation associated with involuting cells just beneath the equator of the early gastrula. When transplanted to the ventral marginal zone (VMZ) of a host embryo, the DMZ will induce a secondary axis complete with head, neural tube, notochord and somites. Hans Spemann and Hilde Mangold (Spemann and Mangold, 1924) first demonstrated this phenomenon using differently pigmented newt gastrulae. By transplanting the DMZ of one newt gastrula to the VMZ of a differently pigmented host, they showed that the resulting embryo formed a second dorsal axis on the ventral side of the embryo. This second axis included all of the dorsal tissue types, such as notochord, somites, and neural tube. Whereas the notochord was always composed of donor cells, much of the neural tube and somites were composed of host cells that would normally have formed ventral tissues such as skin and blood. Spemann and Mangold called the transplanted region the organiser, due to its ability to reorganise surrounding cells that under normal conditions would differentiate into completely different fates. Today the DMZ of *Xenopus* embryos is known as the Spemann organiser, or just organiser, in honour of Spemann's contributions to amphibian embryology.

The organiser serves two signalling functions; firstly it patterns the adjacent marginal zone into tissues of varying dorsal to ventral fates and secondly, it diverts cells of the overlying ectoderm from an epidermal to neural fate (Harland and Gerhart, 1997). A number screens in the early 1990's isolated various organiser specific genes in the hope of identifying the signal, or signals, that were responsible for the organiser phenomenon (Blumberg et al., 1991; Sasai et al., 1994; Smith and Harland, 1992). Surprisingly very few of the genes that were isolated coded for the expected

signalling molecules, most appeared to be inhibitors of signalling molecules found throughout the embryo. All however were able, to varying degrees, mimic the organiser transplant experiments when mis-expressed in the VMZ of *Xenopus* embryos. Three of these, Chordin (Chd), Noggin and Follistatin, inhibited bone morphogenetic protein-4 (BMP-4) (Fainsod et al., 1997; Piccolo et al., 1996; Zimmerman et al., 1996), a TGF- β related signalling molecule first identified as an osteogenic factor (Wozney et al., 1988). A recent study depleting these three BMP inhibitors using antisense morpholino oligonucleotides has once again highlighted their importance for the establishment of dorsal cell fates. Loss of Chd, Noggin and Follistatin from *X. tropicalis* embryos leads to a complete absence of dorsal structures (Khokha et al., 2005). Inhibitors of BMP signalling are also responsible for anterior neural induction during gastrulation in the dorsal ectoderm. By inhibiting BMP-4 they prevent dorsal ectoderm from becoming epidermis and divert it to anterior neural tissue. For an excellent review of all aspects of the previous sections please refer to De Robertis & Sasai (2004).

1.2.4 Patterning the marginal zone - Bone morphogenetic proteins

BMP-4 is fairly uniformly expressed in *Xenopus* blastulae, but is lost from the organiser during gastrulation (Fainsod et al., 1997; Hemmati-Brivanlou and Thomsen, 1995). Injection of *bmp-4* mRNA into the organiser causes a loss of dorsal structures, in both the mesoderm and ectoderm, and in extreme cases injected embryos are completely ventralised (Dale et al., 1992; Jones et al., 1992). This effect is concentration dependent, with high concentrations giving more ventralised phenotypes than low concentrations (Dale et al., 1992; Dosch et al., 1997). Injection of dominant-negative receptors for BMP-4 produced the opposite effect, with both the

mesoderm and ectoderm being dorsalised (Graff et al., 1994; Suzuki, 1995). Once again, the effect is concentration dependent, with high concentrations giving more dorsalised phenotypes than low concentrations (Dosch et al., 1997). Similar concentration-dependent effects have been obtained for both Chd and noggin, presumably through inhibition of BMP-4 signalling (Dosch et al., 1997; Jones and Smith, 1998). Taken together, these results suggest that BMP-4 acts in a concentration dependent manner to specify dorsal-ventral fates, with ventral fates specified by high concentrations and more dorsal fates specified by progressively lower concentrations. This BMP-4 activity gradient is established by a dorsal (high) to ventral (low) gradient of antagonists secreted by the organiser (Dale and Wardle, 1999).

The BMP signalling pathway, and its role in establishing dorsal-ventral cell fates, is also conserved in invertebrates, although due to an inversion of axes during evolution the pathway specifies dorsal cell fates as opposed to ventral in these organisms (De Robertis and Sasai, 1996; Holley et al., 1995). The *Drosophila* orthologue of BMP-4 is Decapentaplegic (Dpp) and ectopic expression of BMP-4 in the *Drosophila* embryo can mimic the effect of Dpp (Padgett et al., 1987; Padgett et al., 1993). High concentrations of Dpp are required to specify the extreme dorsal cell types, such as the amnioserosa, and at low concentrations to specify cells of more ventral characteristics (Ferguson and Anderson, 1992a; Irish and Gelbart, 1987). It is antagonised by short gastrulation (Sog), a protein expressed in the ventral half of the blastoderm that was subsequently shown to be the orthologue of *Xenopus* Chd (Francois and Bier, 1995; Francois et al., 1994). Sog diffuses into the dorsal half of the embryo where it establishes a dorsal (high) to ventral (low) gradient of Dpp

activity (Srinivasan et al., 2002). The establishment of dorsal-ventral cell identities using Chd/Sog and BMP-4/Dpp is therefore a fundamental process conserved throughout evolution (De Robertis and Sasai, 1996; Holley et al., 1995). Consistent with this, genetic studies in *Drosophila* identified additional genes that modulate Dpp activity in the dorsal ectoderm, including *twisted gastrulation (tsg)* and *tolloid (tld)* (Ferguson and Anderson, 1991), both of which have homologues in *Xenopus* that regulate the activity of BMP-4 (De Robertis and Kuroda, 2004).

1.3 The Tolloid Metalloproteases

The Tolloid (Tld) family are Zn^{2+} dependent extra-cellular proteases related to the astacin family of endopeptidases (Bond and Beynon, 1995). They are a highly conserved family that have been found in many species, including humans, mouse, chick, *Xenopus*, zebrafish, sea urchins, molluscs, and *Drosophila*. All have a conserved astacin like protease domain at the N-terminus and an array of C-terminal CUB and EGF-like modules thought to convey different properties to the enzyme. The first mammalian member of the Tld family, bone morphogenetic protein 1 (BMP-1), was purified from demineralised bone extracts, along with BMP-2 and BMP-3 (Wozney et al., 1988). The latter proteins belong to the TGF- β super-family of extracellular signalling molecules that include BMP-4 and display no homology to BMP-1. It was subsequently discovered that BMP-1 is closely related to *Drosophila* Tld (Shimell et al., 1991), which was known to regulate the activity of the Dpp (Ferguson and Anderson, 1992b). Since Dpp is highly homologous to both BMP-2 and BMP-4 in mammals (Padgett et al., 1987) and has similar biological activity (Padgett et al., 1993; Sampath et al., 1993) it was thought that BMP-1 might regulate

TGF- β signalling in mammals. BMP-1 is a shorter protein than *Drosophila* Tld, lacking a number of C-terminal domains (see section 1.4.3). However, a second mammalian protease, with an identical domain structure to that of *Drosophila* Tld, was subsequently identified and shown to be encoded by a splice variant of the same gene that encodes BMP-1 (Janitz et al., 1998; Takahara et al., 1994b). This mammalian gene is now known as *tld* and encodes at least three different splice variants – Tld, BMP-1, and BMP-1-His (which has a histidine rich domain added to the C-terminus of BMP-1). In addition to *tld*, a number of distinct but closely related genes have subsequently been identified in a range of different species. These include *Drosophila tolkin*, which is also known as *tolloid-related* (*tlr*) (Finelli et al., 1995; Nguyen et al., 1994), sea urchin *SpAN/BP10* and BMP-1 (Hwang et al., 1994; Lepage et al., 1992; Reynolds et al., 1992), *Aplysia tld* (Liu et al., 1997), zebrafish *tolloid* (*ztld*) (Blader et al., 1997), *Xenopus xolloid* (*xld*) and *xld-related* (*xlr*) (Dale et al., 2002; Goodman et al., 1998), and mammalian *tld-like 1* and 2 (Scott et al., 1999). With the exception of *SpAN/BP10*, which has a slightly different C-terminal module arrangement, all give rise to proteins that have an identical protein arrangement to that of Tld (Figure 1.1). There are, however, no known splice variants of these genes.

1.4 The Tld protein arrangement

1.4.1 The Astacin protease domain

Astacin was first isolated as a 200 amino acid protease from the stomach of the fresh water crayfish *Astacus astacus* L. (Pfleiderer et al., 1967) and contains a Zn²⁺ binding motif which is common to many Zn²⁺ dependent proteases (Dumermuth et al., 1991; Stocker et al., 1993). The active site binding motif has a consensus sequence of **HExxHxxGxxH**, which forms a α -helix, with the three histidines responsible for co-

ordinating the Zn^{2+} ion. Astacin belongs to a superfamily of Zn^{2+} endopeptidases known as the metzincins (Bode et al., 1994; Bond and Beynon, 1995), which include the Tlds, adamalysins (ADAMs), matrixins (MMPs), and bacterial serralysins. The metzincins are characterised by a unique protein signature known as the "Met-turn" that returns the amino acid chain to the active site, helping to co-ordinate the catalytically important metal ion (Bode et al., 1994) (Figure 1.2). In astacin the Zn^{2+} ion is co-ordinated by three histidine residues (His92, 96, 102), a bridged H_2O molecule, a glutamic acid (Glu93), and a tyrosine residue (Tyr149) (Bode et al., 1992). Glu93 is catalytically important by acting as a proton shuttle in the cleavage reaction and mutation of this residue results in no detectable activity from the enzyme. Tyr149 is believed to be involved in substrate binding and transition state stabilisation, while the H_2O molecule provides a source of protons for the cleavage reaction (Gomis-Ruth et al., 1993; Yiallourous et al., 2000).

1.4.2 The pro-peptide domain and activation of protease activity

Following the signal sequence, which is required for secretion, the Tlds contain a pro-peptide that must be removed to yield an active form of the protease. The pro-peptide sequence is highly variable between different members of the Tld family, but most contain a highly conserved 21 amino acid motif that includes a single unpaired cysteine residue. The role of this cysteine is currently unknown, although it may form a disulphide bond with an as yet unidentified protein. Alternatively, it may function as a "cysteine-switch" acting to regulate the activity of the protease domain. Cysteine switches have been found in other members of the metzincin superfamily, including MMPs and adamalysins such as ADAM-12 (Loechel et al., 1999; Springman et al., 1990). In these proteins a single conserved cysteine in the pro-peptide co-ordinates

Figure 1.1 Domains of the Tld metalloproteases

A schematic diagram outlining the domains making up each member of the Tld family. BMP-1 contains 3 CUB domains with the 2nd and 3rd separated by a single egf-like domain (Yellow box). Tld, Tll-1, Tll-2, Tolkin, zTld, Xld and Xlr all have 5 domains with the 3rd flanked by two egf-like domains. The sea urchin protease SpAN/BP10 contain an astacin domain followed an egf-like domain and two CUB domains separated by a Threonine rich region (Blue box). Signal/pro-peptide, yellow oval; protease domain, Green; CUB domains, Red; egf-like domain, Yellow.

Figure 1.2 The “Met-turn”

A diagram representing the active site of an Astacin protease, including the **HExxH** Zn²⁺ binding motif forming an α -helix. The methionine, which characterises the “Met turn”, is highlighted with a red circle. (after Bode et al., 1994)

Figure 1.1

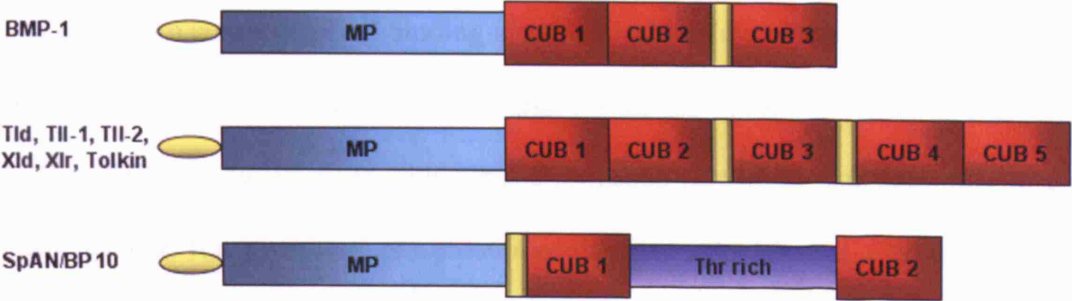
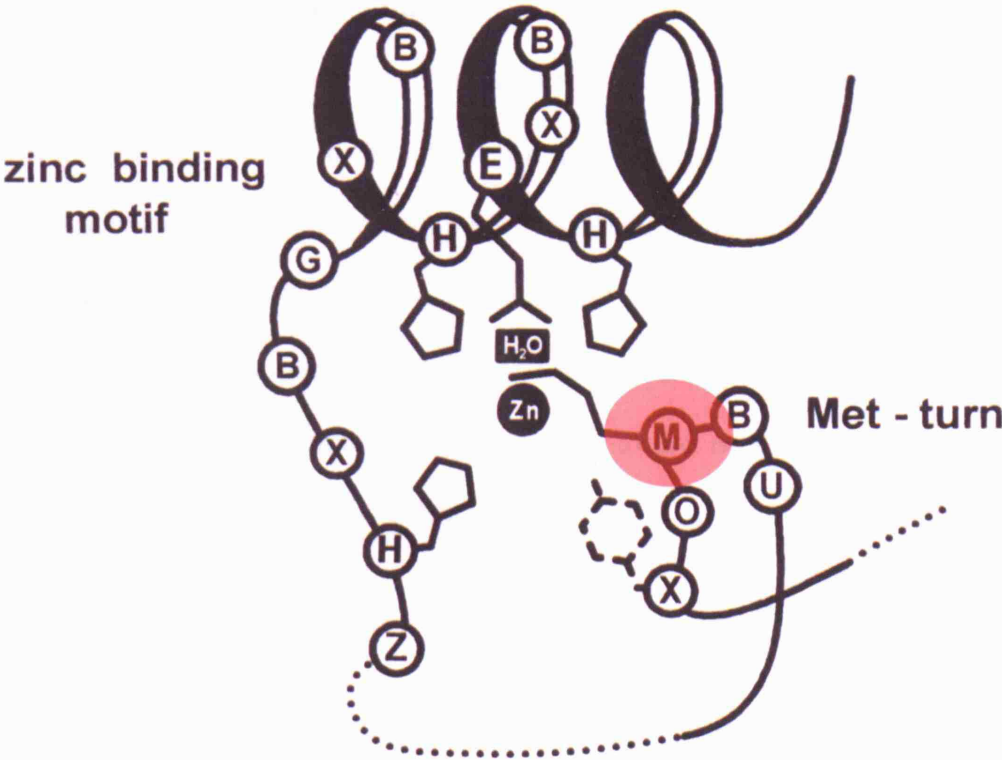


Figure 1.2



the Zn^{2+} ion of the active site, displacing the catalytically important H_2O molecule and preventing substrate access. However, this “cysteine switch” mechanism has yet to be demonstrated for the Tlds. An MMP from mouse, CA-MMP, has been identified that contains a conserved Zn^{2+} binding motif but lacks a “cysteine switch” (Pei, 1999), showing the mechanism does not hold true for all Zn^{2+} proteases. The pro-peptide domain of Tlds always ends with a basic sequence (R-x-x-R) that conforms to the cleavage site for furin-like pro-protein convertases (Seidah and Chretien, 1999). Cleavage at this site is essential for proteolytic activity since the N-terminal alanine of the mature protein is buried deep within the protease, where it forms a hydrogen bond with the active site (Bode et al., 1992). This would not be possible in the presence of the pro-peptide.

1.4.3 CUB and EGF-like domains

The protease domain is followed by an array of CUB and EGF-like (see below) modules that are thought to be involved in protein-protein interactions. In BMP-1, three CUB domains are found with the second and third domain separated by a single EGF-like domain. Tld contains two additional CUB domains with two EGF-like domains flanking the third CUB module (see Figure 1.1). Most other members of the Tld family have an identical CUB arrangement to that of Tld itself, the exceptions being the sea urchin proteins SpAN and BP10.

1.4.3.1 CUB domains

CUB domains are widespread modules first identified in the serum complement serine proteases Cls and Clr, a sea urchin protein called Uegf, and BMP-1. Hence the name CUB was adopted (Complement subcomponents Clr/Cl, Uegf, BMP1) (Bork and

Beckmann, 1993). The module is also found in many other proteins including the procollagen C-proteinase enhancer (PCPE) (Takahara et al., 1994a), mannose binding protein (MBP), serine associated protease-2 (MASP-2) (Feinberg et al., 2003), Uvs-2 (Bork, 1991), and members of the ADAMTS family (Zheng et al., 2001). CUB domains are approximately 110 amino acids long and, with the exception of the serum complement proteins, contain 4 conserved cysteine residues forming two disulphide bridges (Bork and Beckmann, 1993).

Much evidence indicates that CUB domains are involved in protein-protein interactions. For example, MASP-2 consists of two CUB modules separated by a single EGF-like domain and it is this structure that mediates the binding of MBP (Feinberg et al., 2003). PCPE consists of two CUB domains, and a non-homologous domain, that binds both BMP-1 and procollagen type 1 α 1 (proColl1 α 1), enhancing the processing of the collagen into a mature fibril (Adar et al., 1986; Takahara et al., 1994a). Two naturally occurring isoforms of PCPE exist, which retain enhancer activity and consist of only the two CUB domains (Kessler and Adar, 1989), indicating the binding capacity of PCPE resides in these modules. Further evidence that CUB domains are involved in substrate binding has been elucidated by the action of dominant-negative versions of the Tlds. Complete removal of the protease domain from the *Xenopus* Tlds (dnXBMP-1, dnXld) (Wardle et al., 1999), or its inactivation by point mutation within the active site (Xld^{Tyr296-Asn}) (Piccolo et al., 1997), produces a dominant-negative effect. Xld^{Tyr296-Asn} rescues ventralisation caused by over-expression of wild-type Xld and prevents cleavage of the BMP inhibitor chordin (Chd) by the protease. DnXBMP-1 and dnXld also prevent the wild-type proteases from cleaving Chd and can compensate for each other *in vivo* (i.e. dnBMP-1 can

block cleavage of Chd by Xld and vice versa) suggesting some functional redundancy exists between these two proteases and their CUB domains. Additionally, dominant negative Tlds dorsalise ventral mesoderm in *Xenopus* VMZ assays, an effect most likely due to binding Chd and protecting it from the endogenous metalloproteases. Despite indications that the CUB domains of the Tlds are involved in substrate binding, little is understood as to which domains are involved in this process. Hartigan et al., (2003) have recently shown that the 1st CUB domain of BMP-1 is required for secretion of the protease and the 2nd CUB domain for processing proColl α 1.

1.4.3.2 EGF-like domains

EGF-like domains have a similar structure to the extracellular signalling molecule, epidermal growth factor (EGF) and can be separated into two subgroups; calcium binding (cbEGF) and laminin-like EGF domains. The cbEGF-like domains contain a calcium binding sequence at their N-terminus while the Laminin-like domains contain an extra disulphide bond, but no residues for calcium binding (Stenflo et al., 2000). EGF-like domains have been implicated in conveying calcium dependence on CUB domains (Handford et al., 1990). The EGF-like domains of the Tlds are cbEGF domains, although whether their capacity to bind calcium is required for enzymatic activity remains open to question. As we can see from a comparison with the EGF-like domain of MASP-2, a known cbEGF (Feinberg et al., 2003), both Xld and XTld have a **CxCxxGFxxxxxxxxC** motif common to cbEGF domains and a conserved aspartic acid/asparagine hydroxylation site. A β -hydroxylated aspartic acid/asparagine provides one of the calcium binding ligands in MASP-2 (Feinberg et al., 2003), as do aspartic acid, glutamic acid, valine and tyrosine residues. The EGF-

like domains of Xld and XTld appear to have residues classed as conservative changes in each of these positions, making it reasonable to assume that all will bind calcium.

While showing the necessary motifs required for binding calcium, studies on mammalian Tlds suggest that this cation may not be important for their activity. Three recent studies have shown that the EGF-like domains of BMP-1 and mTld are dispensable for cleavage of both proColl α 1 and Chd (Garrigue-Antar et al., 2004; Hartigan et al., 2003; Petropoulou et al., 2005). In addition, removal of the EGF-like domains of mTld actually allow Chd cleavage *in vitro* (Petropoulou et al., 2005). BMP-1, however, has many substrates and it is therefore possible that calcium binding is required for cleavage of at least some of these.

1.5 Cleavage recognition sites

Although the protease domain of Tld is most closely related to astacin, substrate cleavage recognition sites vary between the two enzymes. In astacin the cleavage site can be laid out as an eight amino acid sequence, with cleavage occurring between the fourth (P1) and fifth (P1') residues.

P4 P3 P2 P1 | P1' P2' P3' P4'

A proline is often found in the P2' position, non-polar, hydrophobic residues such as alanine and tyrosine are found in positions P4, P3, P3' and P4' and basic amino acids like lysine and arginine reside in the P2 and P1 positions. In astacin substrates the P1' position is normally occupied by an alanine, threonine, serine, or glycine, but never

aspartic acid (Zwilling, 1994). Yet aspartic acid occupies the P1' position in all the known substrates for members of the Tld family. Table 1.1 provides a summary of the cleavage sites for all known Tld substrates, with the proteolytically important aspartic acid highlighted in red. This conserved aspartic acid is required for C-terminal pro-peptide removal of the $\alpha 2$ chain of procollagen type 1 (proCol1 $\alpha 2$) and removal of a C-terminal fragment of perlecan, its removal renders these proteins resistant to Tld cleavage (Gonzalez et al., 2005; Lee et al., 1990). The presence of the conserved aspartic acid in all of the presently known Tld substrates suggests that this amino acid is important for cleavage of all substrates. Like astacin, non-polar hydrophobic residues are often found in the P4 and/or P3 positions of Tld substrates, most commonly tyrosine or methionine, and in the P3' and P4' positions alanine and proline are commonly found (Table 1.1 bold lettering). However, proline is never found in the P2' position and the basic residues lysine and arginine are rarely found at P2 and P1, as they commonly are in astacin.

1.6 The Tlds in embryonic development

Both Dpp/BMP-4 and Sog/Chd play a major role in the establishment of dorsal-ventral cell fates during gastrulation. *Tld* was first identified in a screen for lethal *Drosophila* mutations that disrupt embryonic patterning and subsequently shown to have a role in dorsal-ventral patterning during the cellular blastoderm stage (Ferguson and Anderson, 1991). The dorsal-ventral axes are patterned during cellular blastoderm stages when cells differentiate according to their dorsal-ventral position within the organism. Cells of the dorsal-midline give rise to extra-embryonic structures such as the amnioserosa, while

Table 1.1 – Cleavage sites of the Tld substrates

Substrate	Cleavage Site
Chd	N-term RSYS DRGE
	C-term PMQA DGPR
proCol1 α 1	YYRA DDAN
proCol1 α 2	FYRA DQP
proCol2 α 1	YMRA DQAA
proCol3 α 1	PYYG DEPM
proCol5 α 1	Yet to be determined
proCol5 α 2	EFTE DQA
proCol11 α 1	Yet to be determined
proCol7 α 1	SYAA DTAG
Lysyl oxidase	RMVG DDPY
Lysyl oxidase-like 1 (Contains two potential cleavage sites)	VAVG DSTG
	VRSS DAPP
Biglycan	FMMN DEEA
Laminin-5 α -3 chain	Yet to be determined
Laminin-5 γ -3 chain	CYSG DENP
Myostatin	NISK DAIR
Osteoglycin	QLQK DET/V
Endorepellin	SGGN DAPG

ventral cells give rise to the mesoderm. In between, ventral-lateral regions differentiate into the neural tube and, along with the dorsal lateral region, the epidermis. In invertebrates however the embryonic axes are reversed, meaning in *Drosophila* dorsal regions are homologous to ventral regions in vertebrate species and vice versa (De Robertis and Sasai, 1996).

A number of additional mutations were identified that also disrupted dorsal-ventral patterning at the same stage, including *sog*, *dpp*, *twisted gastrulation (tsg)*, *screw (scr)* and *zerknüllt (zen)*. The *tld* mutant produced a similar, but weaker, phenotype to that of *dpp*, whereby dorsal structures are lost to an expanding neurogenic ectoderm (Ferguson and Anderson, 1992b; Irish and Gelbart, 1987). Genetic crosses of *tld* and *dpp* mutants suggested that an interaction occurred between the two gene products, and that multiple copies of wild-type *dpp* could compensate for mutations in *tld*. Conversely, multiple copies of wild-type *tld* could not compensate for the mutations in *dpp*. These studies indicated that Tld functioned upstream of Dpp and that Tld was required for maximal signalling by Dpp in the dorsal ectoderm (Childs and O'Connor, 1994; Ferguson and Anderson, 1992b; Finelli et al., 1994; Shimell et al., 1991). The close homology of Tld to BMP-1 and the fact that the latter was first purified along with BMP-2 and BMP-3, both close homologues of Dpp, suggested that the interaction was evolutionarily conserved between vertebrate and invertebrate species (Padgett et al., 1993; Wozney et al., 1988). Observations that *sog* counteracts *dpp* and that Chd binds and inhibits BMP-4 (Francois et al., 1994; Piccolo et al., 1996) led to a model whereby Tld enhanced Dpp signalling by counteracting the effect of Sog (Marques et al., 1997) and similarly, Xld counteracted the BMP-4 inhibition by Chd (Piccolo et al., 1997).

Sog and Chd both consist of four cysteine-rich (CR) domains, three grouped closely in the C-terminal region and a fourth at the N-terminus separated from the others by a non-homologous domain (Sasai et al., 1994). Sog/Chd binds to Dpp/BMP-4 via these CR domains and prevents the signalling molecule from interacting with its receptor (Larrain et al., 2000). Biochemical analysis of *Drosophila* Tld showed it was able to cleave Sog in three places when in the presence of Dpp, in doing so preventing Dpp's attenuation by Sog (Figure 1.3) (Marques et al., 1997). This mechanism was shown to be conserved across the species in which Tld-related metalloproteases have been identified, including Zebrafish, Mouse and *Xenopus* (Blader et al., 1997; Dale et al., 2002; Piccolo et al., 1997; Scott et al., 1999; Wardle et al., 1999), although these proteases do not require the presence of BMP-4 to cleave Chd.

In *Xenopus*, three members of the Tld family have been identified, XBMP-1/XTld, Xld, and Xlr (Dale et al., 2002; Goodman et al., 1998; Piccolo et al., 1997). All but XTld, have been shown to cleave *Xenopus* Chd, producing fragments with reduced affinity for BMP-4. In *Drosophila*, the opposing expression of *tld* to *sog* suggests that Sog establishes a gradient of Dpp activity and indeed a gradient of Sog protein can be visualised in the blastoderm controlled by Tld produced in the dorsal half of the embryo (Marques et al., 1997; Srinivasan et al., 2002). In *Xenopus* however XBMP-1 and Xld are present throughout the marginal zone, suggesting a more complicated mechanism

Figure 1.3 Tld cleaves Chd/Sog releasing bound BMP-4/Dpp

Members of the Tld family cleave Chd/Sog in two places, once next to the N-terminal CR1 domain and again between the CR3 and CR4 domains near the C-terminus. This produces fragments with reduced affinity for BMP-4/Dpp, consequently releasing the bound molecule and allowing it to function.

Figure 1.4 Assembly of mature collagen fibrils in the ECM

Procollagen fibres are secreted to the ECM with N and C-terminal pro-peptides. These are removed by specific N- and C- proteinases, releasing an insoluble collagenous domain. These assemble in to mature fibrils of various diameters mediated by lox.

Figure 1.3

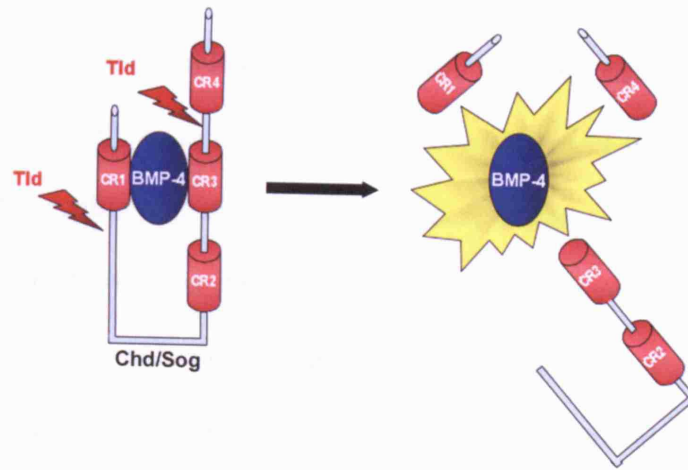
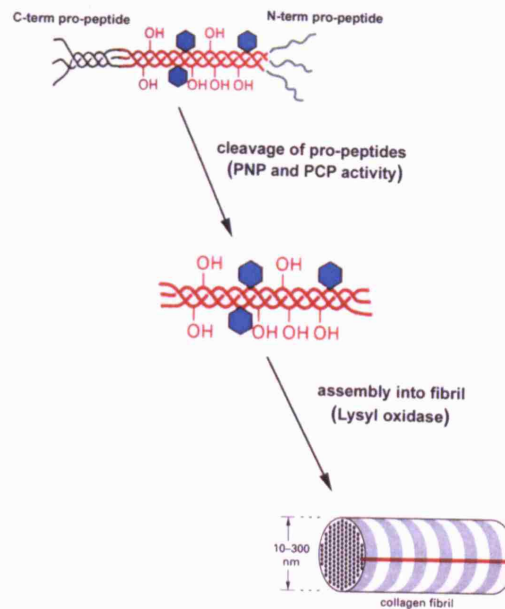


Figure 1.4



may be in operation. While the model of a simple BMP-4 gradient, established by secretion of Chd from the organiser, is often described for the marginal zone, the presence of XBMP-1 and Xld in the dorsal marginal zone suggests otherwise. An alternative explanation is that the role of Tlds in *Xenopus* gastrulation is to restrict the inhibitory ability of Chd to the DMZ and that another mechanism is used to establish the BMP-4 signalling gradient. For instance, noggin, follistatin and Xnr-3 are inhibitors of BMP-4, as is the multifunctional inhibitor cerberus, and all, just like Chd, are secreted molecules expressed in the organiser (Fainsod et al., 1997; Haramoto et al., 2004; Piccolo et al., 1999; Smith and Harland, 1992). Binding data for noggin indicates that it binds BMP-4 with a greater affinity than Chd, but this protein is not inactivated by Tlds. The same can be said for Xnr-3, whose N-terminal pro-peptide inhibits BMP-4 and induces secondary axes in the VMZ (Haramoto et al., 2004). While no diffusion data exists for either noggin or follistatin, a nodal-related family member, Xnr-2, has been shown to diffuse in the manner of a morphogen suggesting Xnr-3 could do the same, thereby establishing a gradient of BMP activity (Williams et al., 2004).

Recent work has also investigated the role of Tsg in the establishment of dorsal-ventral patterns in both *Drosophila* and *Xenopus*. Tsg is a secreted protein that can bind both BMP-4 and Chd (Mason et al., 1994; Oelgeschlager et al., 2000). In *Xenopus*, *tsg* is expressed in the VMZ and there has been conflicting evidence for its role, in some cases acting as an enhancer of BMP-4 and in others attenuating BMP signalling (Chang et al., 2001; Ross et al., 2001; Scott et al., 2001). It is now thought that, in the presence of Tld-like proteases, Tsg enhances the processing of Chd and releases more BMP-4, but in their absence functions to attenuate signalling (Larrain et

al., 2001). Additionally, once the bound BMP molecule has been released, Tsg can serve to accelerate the cleavage of Chd by the Tlds. Scott et al, (2001) suggest that in the presence of Tsg, Chd is in fact cleaved in a third position, immediately N-terminal of the second CR domain. Tsg appears to add a further layer of control to the establishment of a BMP activity gradient and may give rise to fragments of Chd that have a different biological activity and properties in the embryo to those artificially observed. A number of different papers have reported that individual CR domains of Sog or Chd have potent BMP inhibitory properties (Larrain et al., 2000; Yu et al., 2000). Paradoxically, similar fragments are produced by the action of the Tlds on Sog/Chd and Tsg might therefore also be a mechanism by which this effect can be controlled.

1.7 The Tlds in Collagen formation

In addition to their role of regulating Chd in the marginal zone, the Tlds also play an important role in the maturation of fibrillar collagens. Collagens form a major part of the ECM and are categorised into two groups, known as the fibrillar and non-fibrillar. To date 27 different types of collagen have been identified of which seven, Types I, II, III, V, XI, XXIV and XXVII, belong to the fibrillar class (Boot-Handford et al., 2003; Ottani et al., 2002; Pace et al., 2003).

Synthesised in fibroblasts, individual collagen fibres are comprised of three different domains, including a central long collagenous domain made up of Gly-X-X repeats forming a tight right-handed helix. A proline residue frequently occupies the X positions of this repeat. Flanking this central domain are both C- and N-terminal non-collagenous (NC1 and NC2 respectively) globular pro-peptides. Prior to secretion

individual chains combine to form a homo- or heterotrimeric left-handed triple helix, in part mediated by the NC1 domain. The NC1 regions of three collagen fibres interact and the collagenous domains wind around each other to form a left handed chain with the two globular NC domains at either end. In this state the Procollagen is secreted to the ECM as a soluble molecule. In the ECM the NC1 and NC2 domains are removed by procollagen C- (PCP) and N-proteinases (PNP) respectively, where by the insoluble collagenous domains of separate molecules condense with each other to form a mature fibril (Figure 1.4) (Boot-Handford and Tuckwell, 2003; Ottani et al., 2002; Prockop and Hulmes, 1994; Prockop and Kivirikko, 1995). This condensation is mediated by lysyl oxidase (Lox), which oxidises lysine residues along the collagenous domain allowing spontaneous condensation of mature collagen fibres (Csiszar, 2001; Kagan, 1986).

The fibrillar collagens themselves are also divided into two groups, generally based on their relative abundances in tissues. The more abundant, major group consists of collagens type I, II and III, while the remaining four are termed as minor collagens and found much less abundantly. All seven fibrillar collagens are formed by trimeric combinations of 12 different polypeptide chains; $\alpha 1$ and $\alpha 2$ (I), $\alpha 1$ (II), $\alpha 1$ (III), $\alpha 1$, $\alpha 2$ and $\alpha 3$ (V), $\alpha 1$, $\alpha 2$ and $\alpha 3$ (XI), $\alpha 1$ (XXIV) and $\alpha 1$ (XXVII). These take the form of homotrimeric (e.g. $\alpha 1$ (II)₃) or mixed heterotrimeric (e.g. $\alpha 1$ (V)₂ $\alpha 2$ (V)₁) combinations. This depends on the tissue or collagen type required (Boot-Handford and Tuckwell, 2003; Ottani et al., 2002; Pace et al., 2003).

The minor fibrillar collagens play a role in controlling the formation and diameter of collagen fibres predominantly made of the major chains. For example, collagen type

V is thought to control the formation of collagen type I fibrils. Mouse null mutants for *proCol5a1* develop normally but die at embryonic day 10 (E10) due to cardiovascular defects (Wenstrup et al., 2004). Analysis of type I collagen in these knock out embryos revealed that they were abnormally large in diameter and poorly formed. Collagen type XI, on the other hand, associates with type II collagen and disrupting the ratio of these two fibres results in an abnormal fibre diameter (Fernandes et al., 2003; Garofalo et al., 1993; Wu and Eyre, 1995).

The PCP responsible for the removing the C-terminal pro-peptide of fibrillar collagens, was shown to be identical to BMP-1 and Tld (Kessler et al., 1996; Li et al., 1996). BMP-1/PCP removes the C-terminal propeptide from procollagens type I, II and III (Hojima et al., 1985; Prockop and Hulmes, 1994), both the N and C terminal propeptides from procollagen type V (Imamura et al., 1998; Kessler et al., 2001) and the N-terminal propeptide from procollagen type XI (Medeck et al., 2003). This activity can be enhanced by the presence of the PCP enhancer (PCPE), which binds both procollagens and BMP-1/PCP (Adar et al., 1986; Hulmes et al., 1997; Kessler and Adar, 1989).

Interestingly Lox, which mediates the cross-linking of PCP processed collagen fibres, is also activated by the Tlds (Panchenko et al., 1996; Uzel et al., 2001). Removal of an N-terminal pro-peptide activates the Lox catalytic domain, which can then deaminate lysine residues along individual collagen molecules allowing them to spontaneously condense into mature fibres (Csiszar, 2001; Kagan, 1986). A number of other proteins closely related to Lox, and sharing a highly conserved LOX catalytic domain, have been identified. At present it is unknown if they too play a role in

collagen formation, but mammalian Lox-like 1 (Loxl-1) is also cleaved in its pro-region by BMP-1 (Borel et al., 2001).

Two other factors implicated in collagen fibrillogenesis have also been described as substrates of BMP-1. Biglycan (Bgn) is a small leucine rich proteoglycan (SLRP) closely related to decorin, a protein that interacts with and influences the distance between type I collagen fibrils. Bgn also interacts with this fibre and is processed by BMP-1 (Schonherr et al., 1995; Scott et al., 2000). Additionally, an SLRP known as osteoglycin (Ogn) enhances collagen fibrillogenesis, although the mechanism by which this occurs is unknown, and is dependent on processing by BMP-1 (Ge et al., 2004). Ogn is also processed by both mTld and mTll-1.

Although playing a role in the solubility and subsequent fibril formation, recent observations may indicate a further role of the pro-peptides. The N-terminal domain of a proCol2 α 1 splice variant contains three 69 amino acid CR domains similar to those found in Chd and, like Chd, have anti-BMP activity when injected into *Xenopus* embryos (Larrain et al., 2000; Sandell et al., 1991). Whether this is as an important factor in patterning the embryo remains to be seen, but proCol2 α 1 is expressed in dorsal regions of *Xenopus* neurulae (Su et al., 1991), when inhibition of BMP-4 is thought to be crucial for the development of dorsal tissues.

1.8 Other substrates of the Tlds

BMP-1 has been implicated in the cleavage of a number of other components of the ECM and also the activation of a latent member of the TGF- β family. Laminins are major components of the basal lamina, consisting of an α , β and γ subunit (Colognato

and Yurchenco, 2000). After secretion to the ECM, the γ -2 and α -3 chains of Laminin-5 undergo proteolytic processing. The γ -2 chain is processed by BMP-1 and it is thought that the α -3 chain may be cleaved in a similar fashion (Amano et al., 2000; Veitch et al., 2003). The processing of γ -2 occurs in the intervening sequence between the third and fourth laminin G domains (LG) at the C-terminus of the protein. In doing so it regulates the interaction of the LG domains with different components of the basal lamina, including dystroglycans and syndecans. Another protein, the heparin sulphate proteoglycan perlecan, also contains a series of LG-like repeats at its C-terminus. The C-terminal domain of perlecan is removed by Tld proteases and this fragment, known as endorepellin, possesses its own biological activity acting on endothelial cells (Gonzalez et al., 2005). Finally, a TGF- β family member, myostatin, is regulated by the Tlds in mammals. Myostatin is a negative regulator of skeletal muscle and found in the blood of mice as an inactive precursor. All four Tlds found in mammals (BMP-1, mTld, mTll-1 and Tll-2) remove the pro-peptide rendering the growth factor active, although mTld is the least efficient at this (Wolfman et al., 2003).

1.9 Summary and Aims

The Tlds contribute to an evolutionarily conserved system of patterning cell fates along the dorsal-ventral axis of early embryos, in addition to playing a central role in the formation of mature collagen fibres in the ECM. In *Xenopus* embryos studies have mainly focussed on the Tlds role in regulating Chd inhibition of BMP signalling in the marginal zone. Whereas research in mice and chicks has centred on their function in ECM formation and maturation. One aim of this thesis is to extend the studies of *Xenopus* Tlds into ECM formation by identifying genes that encode

orthologs of proteins already characterised as substrates in other organisms. A second aim of this thesis is to further characterise the roles played by different domains within the Tld metalloprotease, Xld. More specifically, how these domains interact with the BMP inhibitor Chd. Finally, an inherent problem of proteases with multiple substrates is how are different substrates identified and the activity of the enzyme controlled. To this end I have attempted to identify an inhibitor of Tld function, which would allow further control of substrate cleavage in the early embryo.

Chapter 2

Materials and Methods

2.1 Materials and Solutions

All chemicals were obtained from Sigma, BDH or Fisher unless otherwise stated. Molecular biology reagents were obtained from Roche Biomolecular or Promega unless otherwise stated and PCR primers purchased from Sigma-Genosys. Autoclaved solutions indicated in the text were done so at 121°C for 30 minutes.

2.1.1 DNA Constructs and PCR primers

PCR primers were obtained from Sigma-Genosys. All ESTs were identified using the Genbank EST databases and obtained from either HGMP or NIBB (Japan). Table 2.1B summarises the ESTs, their Accession and IMAGE/NIBB clone numbers and the RT-PCR primers used in this Thesis. Other constructs used have been identified as such in the text. Primers used for Xld CUB deletions are found in Appendix G.

Table 2.1A – General RT-PCR primers

Gene	RT-PCR primers
<i>Ornithine decarboxylase</i> (Agius et al., 2000)	Fwd: cagctagctgtggtgtgg Rev: caacatggaaactcacacc
<i>Xenopus procol1α1</i> (Goto et al., 2000)	Fwd: atcttcagggtcacggcaggt Rev: attcaacggaccctctggac
<i>Xenopus procol2α1</i> (Sasai et al., 1996)	Fwd: aggcttggtggtcctcaagg Rev: tgtaacgcatagggtcgggtcc

All PCR primers were made to a concentration of 1 μ g/ μ l and diluted 1:4 to give a working concentration of approximately 50 pmol/ μ l.

Table 2.1B – ESTs and RT-PCR primers used in this thesis

EST	Accession number	Clone	RT-PCR primers
<i>Xlox</i>	AJ919279	4680378 (IMAGE)	Fwd: tggatgccagcagtcacagaagag Rev: ccatgtatcgtaaatacacaacgc
<i>Xloxl-1</i>	AJ919280	4173233 (IMAGE)	Fwd: tgggagagggcttcctgattg Rev: catccatactgtgatagtgtggtg
<i>Xloxl-3</i>	AM040262	XL051k10 (NIBB)	Fwd: caaggaatcacagtaggatgtctgg Rev: acgcttatttgcttcttcgctg
<i>ProCol3α1</i>	BM179481	4964353 (IMAGE)	Fwd: ccaggcatacaactcaac Rev: ttctacctcatcaaaactccc
<i>ProCol5α1</i>	BX844650	5541989 (IMAGE)	Fwd: aggggtgatgatggagagatg Rev: cctgatggctcttcttacc
<i>ProCol5α2</i>	BM179681	4959113 (IMAGE)	Fwd: tgaagtgatggcgcaatc Rev: gatgggcgaaaacatacg
<i>ProCol11α1</i>	BG579352	4203897 (IMAGE)	Fwd: cagtgagttcaagagaggaaag Rev: aaggctctgatgtatgggttg
<i>Xbgn-EST</i>	BG729865	3580939 (IMAGE)	Fwd: gaccgattggctatccagtttg Rev: attgtgcgagtgtgggggttac
<i>laminin α-like</i>	BG892766	3475590 (IMAGE)	Fwd: tgctttccatccaacttcaag Rev: tctccctttttctgtatgtgcc

All ESTs were provided in either pCMV-SPORT6 or pRN3 and where appropriate transferred into pCS²⁺ as indicated in the text.

2.1.2 Solutions and media

***Xenopus* and Embryo solutions**

Hormones

Lyophilised Folligon (Intervet), an FSH substitute, was reconstituted in the resuspension solution provided, separated into 100 i.u. aliquots and stored at -20°C for future use. Lyophilised Chorulon (Intervet), an HCG substitute, was also reconstituted with the solution provided and used immediately. Both lyophilised hormones were stored at 4°C until use.

Embryo culture medium

Normal Amphibian media (NAM) pH 7.5

10 x Stock	1100 mM	NaCl
	20 mM	KCl
	10 mM	Ca(NO ₃) ₂ ·4H ₂ O
	1 mM	EDTA

This stock was autoclaved and stored at RT.

Working solutions of NAM were 50% strength (NAM/2) for embryo manipulations, and 1/20th strength (NAM/20) for general culture. The pH was adjusted by the addition of 2mM HEPES at pH 7.5. Gentamycin antibiotic was added in a concentration of 0.025 mg/ml. For NAM/2 sodium bicarbonate was added.

Modified Barths Solution (MBS)

25 x MBS-A	88 mM	NaCl	25 x MBS-B	0.3 mM	Ca(NO ₃) ₂ ·4H ₂ O
pH 7.6	1mM	KCl		0.4 mM	CaCl ₂ ·2H ₂ O
	2.5 mM	NaHCO ₃		0.8 mM	MgSO ₄ ·7H ₂ O
	10 mM	HEPES-NaOH			

Working solutions for MBS were 1 x MBS for embryo manipulations, and 1/10th MBS for general culture. Both stock solutions were stored at RT.

Embryo/Oocyte Homogenisation buffer

20mM	Tris-Hcl (pH 7.6)
100mM	NaCl
1%	Triton X-100

1 protease inhibitor cocktail tablet (Roche) was dissolved in 10ml of this buffer for homogenisation of oocytes and embryos.

Buffers for Western Blotting

2 x Sample Buffer

1ml	1M Tris-HCl (pH 6.8)
1.6ml	Glycerol
3.2ml	10% SDS

Transfer Buffer

2.91g	Tris
1.465g	Glycine
1.875ml	10% SDS

0.4ml	0.05% (w/v) Bromophenol blue	100ml	Methanol
9.0ml	dH ₂ O	to 500ml with dH ₂ O	
0.8ml	β-mercaptoethanol		

β-mercaptoethanol this was eliminated in for non-reducing buffers.

10 x TBS Stock

250mM Tris-HCl (pH 8.2)

1440mM NaCl

For 1 x TBS this was diluted 10 x in dH₂O. For TBST, 0.1% Tween-20 was added

Whole mount *in situ* hybridisation

MEMFA fixative

0.1M MOPS pH 7.4

2mM EGTA

1mM Magnesium sulphate

3.7% Formaldehyde

Hybridisation buffer (Stored at –20°C)

50% Formamide

5 x SSC pH 7.0 (20 x SSC; 3M NaCl, 0.3M Tri-Sodium citrate, autoclaved)

5 mg/ml Torula RNA (Stock at 25mg/ml in 50% Formamide)

100μl/ml Heparin

1 x Denhardtts solution (From 50x stock)

0.1% Tween-20

0.1% CHAPS

0.25 mM EDTA

(50 x Denhardtts solution; 1% w/v BSA Pentax Fraction V, 1% w/v Ficoll, 1% w/v Polyvinylpyrrolidone)

Phosphate buffered Saline (PBS/PBSAT)

10 PBS tablets (Oxoid) were dissolved 100 ml H₂O and autoclaved, 10 x stock was diluted with nfH₂O to a 1 x PBS solution. For PBSAT, Tween-20 was added at a final concentration of 0.1% to 1 x PBS. Both stored at RT.

Maelic Acid Buffer (MAB)

100 mM Maelic Acid

150 mM NaCl

MAB was autoclaved, and Tween-20 was added at 0.1% when the solution was cool. Stored at RT.

20XSSC

3M NaCl

0.3M Tri-sodium citrate

pH adjusted to 7.0 with NaOH, Autoclaved and stored at RT.

Alkaline Phosphatase Buffer

100mM Tris-Hcl pH 9.5

100mM NaCl

50mM MgCl₂

0.1% Tween-20

5mM Levamisole

Made fresh on each occasion.

Bacterial culture**LB media**

4g Typtone (Invitrogen)

4g NaCl

2g Yeast Extract (Invitrogen)

400ml dH₂O

LB was autoclaved, allowed to cool to ~50°C and filter ampicillin or kanamycin antibiotic added at the appropriate concentration depending on requirement.

LB agar plates

LB was made as above, with the addition of 1.5% Agar, and autoclaved. Ampicillin was added at ~50°C and the agar poured into 90mm Petri dishes. Plates were left to set, and hardened overnight at 37°C.

Solutions for Histology

Romeis fixative

25ml saturated HgCl₂

20ml 5% TCA

15ml 37% Formaldehyde

Borax-Carmine

8g Carmine

10g Sodium Borate

250ml dH₂O

Picoro-Blue black

5g Picric acid

390ml dH₂O

9.75g Naphthol blue black

2.2 Methods

2.2.1 Xenopus techniques

Eggs were obtained from mature female *Xenopus laevis* by subcutaneous injection of 750 i.u. human chorionic gonadotrophin (HCG) (Chroulon, Intervet) into the dorsal lymph sac of the frog approximately 12-16 hrs. prior to requirement. In some cases frogs were primed with 50 i.u. of follicle stimulating hormone (FSH) (Folligon, Intervet) 3 to 5 days before injection of HCG to help induce laying. Eggs were squeezed from a frog induced with HCG into a petri dish and fertilised with a piece of macerated testis dissected from a freshly sacrificed male *Xenopus*. Males were sacrificed by placement for 30-60 minutes in 0.5% MS-222 (Sigma), an anaesthetic which permeates the skin of the frog and at this concentration is lethal. The testis was spread around the eggs, before being made into a suspension with tap water, which was also spread around the eggs, the tap water aids the motility of the sperm. Remaining testis was placed in Foetal Bovine serum (FBS) (GibcoBRL) and stored at 4 °C, for up to 1 week. The eggs were left for approximately 5 to 10 minutes at room temperature and then flooded with fresh NAM/20 or MBS/10 solution. Embryos were left to rotate, then de-jellied in 2% cysteine hydrochloride (Sigma) (pH 8.2) for

approximately 5 to 10 minutes in a large beaker until the jelly coat had dissolved. The embryos were then washed with copious amounts of tap water until all cysteine was removed. Embryos were returned to a fresh Petri dish and left to cultured in either NAM/20 or MBS/10. Dead embryos were removed and discarded.

Embryos were staged using the criteria of Nieuwkoop and Faber (1969). In some cases embryos were also scored for dorsal or ventralisation, this was done using the Dorso-Anterior Index (DAI) developed by Kao and Elinson (1988). Under the DAI normal embryos are scored 5, ventralised embryos between 4 and 0 depending on the severity, similarly dorsalised embryos scored between 6 and 10 again depending on severity of the defect. A detailed description of the criteria used in this study is given in Appendix G.

2.2.2 Embryo manipulations

2.2.2.1 Animal cap removal.

Embryos were cultured in NAM/2 until Stage 8/9. The vitelline membrane of an embryo was removed with sterile No. 5 watchmakers forceps. Using the same forceps two incisions were made opposite each other at the boundary of the animal pole region, the embryo was then rotated 90°, and two similar sized incisions were made in the same fashion. The animal cap was removed from the embryo and left to heal in NAM/2 at RT, or used immediately in a further manipulation. Activin induction of Animal caps was achieved by the addition of diluted activin protein (Sigma).

2.2.2.2 DMZ and VMZ removal.

Embryos were cultured in NAM/2 until the blastoporal lip had formed. Using watchmakers forceps the vitelline membrane was removed and with a tungsten needle two incisions were made either side of the lip about 45° apart. The embryo was then rotated 90° and a further cut made below the lip on the vegetal pole to release the DMZ from the embryo. Excess endodermal cells were trimmed away, and the fragment was left to culture in NAM/2, or used directly in a further manipulation.

Removal of the VMZ involved an identical procedure however a 45° fragment of the marginal zone was removed from directly opposite the blastoporal lip. Again excess

tissue from the vegetal pole was removed, and the fragment left to heal and culture in NAM/2.

2.2.2.3 Oocyte collection

A mature female *Xenopus* was sacrificed in 0.5% MS-222 (Sigma) and the ovarian tissue removed. The ovaries were placed in 2 mg/ml collagenase (Sigma) for 1-2 hours, then placed in 1 X MBS. Mature oocytes were selected and incubated overnight at 16 to 18°C in 1 X MBS. Surviving oocytes were selected the following day for injection.

2.2.2.4 Microinjection of mRNA

De-jellied embryos were placed in agarose coated petri dishes with NAM/20 containing 2% Ficoll (Sigma). Ficoll removes the liquid from between the vitelline membrane and the embryo, and thickens the media which stabilises the embryos during injection. Micro-pipettes were made from 1.0 mm od x 0.58 mm id x 10 cm Borosilicate glass capillaries (Intrafil), pulled to the correct extent using a Narishige PC-10 pipette puller. The tip was cut off using watchmaker's forceps, and the pipette attached to a PLI-100 micro-injection rig (Medical Systems), attached to an O₂ free N₂ gas cylinder. The pipette was back filled with synthetic capped mRNA placed on Nescofilm. The pipette was calibrated by injecting a small amount of mRNA into mineral oil and measuring the size of the bubble produced using an eye-piece graticule attached to the microscope. Injection pressure and time were adjusted until the required volume was obtained. Synthesis of synthetic capped mRNA is described later in this chapter. Embryos were cultured and injected at either the two or four cell stage depending on requirement. For some experiments embryos were cultured until later stages, which is indicated in the text. After injection the embryos were left to culture at 13, 16, or 18°C in NAM/20 + 2% Ficoll until they had passed gastrulation whereby they were transferred to NAM/2 which induces better survival. Embryos placed in NAM/2 before gastrulation will exogastrulate.

2.2.3 DNA/RNA Techniques

2.2.3.1 Agarose gels

DNA or RNA fragments were visualised on 0.5-1.5% Agarose (Sigma) gels containing a small amount of ethidium bromide made with either 1xTBE, or 1xTAE. Buffer type and gel percentage depended on the type and size of fragment to be analysed.

2.2.3.2 Restriction digests

Vector or mini prep DNA was cut using restriction endonucleases (Roche) in 1 x reaction buffer (provided with the enzyme) at 37°C. The length of incubation and amount of enzyme added was dependent on the amount of DNA to be cut.

2.2.3.3 Cleaning of digests

Cleaved vector DNA or fragments to be cloned into another vector were cleaned using the following method. The digest was made to 100 µl with nfH_2O and 1 µl of 100 mg/ml Proteinase K added, this was incubated at 50°C for 30 minutes. An equal volume of phenol:chloroform:isopropanol (25:24:1) was added, and the tube vortexed for 15 seconds. The tube was micro-centrifuged for 12 minutes at 13000 rpm and the upper aqueous phase removed to a fresh tube. This step was repeated, and followed by the addition of an equal volume of chloroform. This was again vortexed, spun for 12 minutes and the aqueous phase removed to a fresh tube. Three chloroform extractions were performed, after the final step the DNA was ethanol precipitated with 7 µl ammonium acetate, and 500 µl 100% ethanol. The DNA was pelleted in a micro-centrifuge spun at 13000 rpm for 10 minutes, then washed in 70% ethanol. The DNA was re-spun for 8 minutes and the pellet left to air dry. The pellet was resuspended in nfH_2O , and checked on an agarose gel.

2.2.3.4 Synthetic mRNA transcription

Synthetic capped mRNA for microinjection was synthesised using either the T3/T7 or Sp6 mRNA *in vitro* transcription kit (Ambion). 20 µg of vector DNA was linearised by restriction digest at 37°C overnight, cleaned by Phenol:Chloroform extraction, ethanol precipitated, and resuspended in nfH_2O to a concentration of 1 µg/µl. The

reaction was set up on ice in a microcentrifuge tube, according to the manufacturers instructions. The reaction consisted of the following.

1 μ l	Vector DNA @ 1 μ g/ μ l
2 μ l	10 x Reaction buffer
10 μ l	2 x NTP/CAP mix
2 μ l	Enzyme mix

Made to 20 μ l with nfH₂O

The reaction was mixed, spun down, and placed at 37°C for 2 hours. 1 μ l of RQ DNase1 was added to the reaction and incubated once again at 37°C for 15 minutes to remove the DNA template. 115 μ l of nfH₂O and 15 μ l of Ammonium acetate stop solution were added. The mRNA was either Phenol:Chloroform extracted and precipitated with Isopropanol or cleaned using RNeasy spin columns (Qiagen) and finally quantified using a spectrophotometer.

2.2.2.5 RT-PCR

RNA isolation

Embryos were collected for RT-PCR and frozen at -80°C until required in 1.5 ml microcentrifuge tubes. Total RNA was isolated using TRIZOL (Invitrogen). Embryos were macerated with a pipette tip, and 1 ml of TRIZOL added to the tube. The mixture was pipetted up and down, and then spun for 10 minutes at 13000 rpm in a cooled microcentrifuge. The supernatant was removed to a fresh tube and left for 5 minutes at RT. 200 μ l of Chloroform was added to the supernatant and vortexed, this was left at RT for a further 3 minutes, and then centrifuged for 15 minutes at 13000 rpm. The aqueous phase was removed to a fresh tube, 500 μ l of isopropanol was added, and then left at RT for 10 minutes to precipitate the RNA. The tube was centrifuged for 10 minutes at 13000 rpm to pellet the RNA. This was then washed in 70% ethanol, and spun for a further 5 minutes at 7500 rpm. After removal of the ethanol the pellet was air dried and resuspended in 100 μ l of nfH₂O.

DNase treatment

To the isolated RNA, 10 µl of RQ DNase1 and 10 µl of 10 x buffer (Promega) was added and incubated at 37°C for 30 minutes. This step removes any remaining genomic DNA that may contaminate the PCR reaction, and give false results. The RNA was subjected to a Phenol:Chloroform and Chloroform extraction, ethanol precipitated and resuspended in 100 µl of nfH_2O .

Reverse transcription of cDNA.

Complementary DNA (cDNA) was synthesised from isolated RNA using either Oligo dT or random primers (Promega) in the following reaction. Oligo dT primers attach to the poly uracil sequence of tRNA transcripts which are reverse transcribed into the poly adenine tail of a cDNA sequence, random primers are able to attach anywhere to the tRNA sequence. Both provide initial points for reverse transcriptase to initiate cDNA synthesis.

dNTP mix (Promega)	2 µl
Oligo dT/Random primers	1 µl
5 x M-MLV RT reaction buffer	10 µl
RNA	5 µl
nfH_2O to a final volume of 43 µl.	

The reaction, assembled on ice, was placed at 65°C for 5 minutes to denature any secondary structure of the RNA, then cooled on ice for 5 minutes. To the reaction 2 µl of M-MLV reverse transcriptase (Promega) was added and then placed at 42°C for 30 minutes. Aliquots of this reaction mix was used directly in the PCR reaction.

PCR reaction

A table of the PCR primers sequences and the conditions used are shown below.

Each PCR reaction contained the following

nfH_2O	18.50 µl
dNTP mix (Promega)	0.5 µl
Fwd. Primer	0.5 µl
Rev. Primer	0.5 µl

Taq Polymerase (Eppendorf)	0.25 µl
10 x reaction buffer	2.5 µl
[³² P]-dCTP	0.25 µl
cDNA	2.0 µl

As a control RNA from Stage 24/25 embryos that had not been reverse transcribed was used. The PCR reaction conditions were as follows

Initial denaturation	94°C	2 minutes
	94°C	20 seconds
Cycles	$x^{\circ}\text{C}$	10 seconds
	65°C	30 seconds
Termination	65°C	5 minutes

x refers to the temperatures specific to each primer pair, individually determined for each primer pair using an Eppendorf Gradient MasterCycler.

5 µl of 5 x TBE sample buffer (Invitrogen) was added to each reaction, and 6 µl loaded onto a 6% TBE precast gel (Invitrogen). The gel was run at 200V for 45-60 minutes, and fixed in 40% methanol/10% acetic acid for 30 minutes. The gel was dried onto Whatman paper at 80°C for 2 hours on a gel drier, and exposed to Fujifilm Super RX medical X-ray film overnight at -80°C.

2.2.2.6 PCR for deletion constructs

PCR primers were designed in opposing directions so that extension would proceed around the whole of the template plasmid. Each reaction contained the following.

nfH ₂ O	14.5µl	
10xBuffer	2.5µl	
dNTP	1.0µl	
Fwd primer	2.0µl	(refer to Appendix H)
Rev primer	2.0µl	(refer to Appendix H)
Pfu Polymerase	2µl	
Template (5ng/µl)	1µl	

The reaction conditions were as follows

Initial denaturation	94°C	2 minutes
----------------------	------	-----------

	94°C	30 seconds
Cycles (x 16)	$x^{\circ}\text{C}$	30 seconds
	72°C	18 minutes
Termination	72°C	5 minutes

x refers to the temperatures specific to each primer pair, individually determined for each primer pair using an Eppendorf Gradient MasterCycler.

The PCR reaction was treated with Dpn1 for 3 - 4 hours to remove the template DNA. Dpn1 only recognises methylated DNA, which is produced by bacteria, thus removing the template and leaving the PCR product. The reaction was then ethanol precipitated, resuspended in 40 μl nfH_2O , and treated with Polynucleotide Kinase-1 (PNK-1) to phosphorylate the blunt ends of the PCR product. After Phenol:Chloroform, Chloroform treatment and ethanol precipitation 1 μg of PCR product was ligated using T4 DNA ligase, with 10mM ATP in a volume of 50 μl . This reaction was performed at 13°C over 2 to 3 days. 25 μl of the ligation was transformed into DH5 α competent cells and plated onto Agar plates containing Ampicillin, and placed at 37°C overnight. Colonies were picked, transfected into 5ml of LB medium containing ampicillin, and again cultured overnight at 37°C. DNA from the overnight cultures was isolated using a QiaQuick mini-prep spin columns (QIAGEN), and analysed by restriction digests.

2.2.2.7 Sequencing

DNA constructs and EST clones were sent to ABC sequencing at Imperial College, London or to Stuart Martin, Sequencing & Genotyping Facility, Dept. of Physiology, UCL. 1 μg of DNA with 12.8 pmol of primer was provided. Sequencing results were provided in *.sea format, edited using EditView, analysed using MacVector and assembled in AssemblyLign (Accelrys).

2.2.4 Protein Analysis.

2.2.4.1 Cell culture and transfection

Approximately 2.8×10^6 HEK293T cells were transferred to 75 cm^3 flasks contain 25 ml Dulbecco's Modified Eagles Medium (DMEM) with 10% Foetal Bovine Serum (FBS) and incubated at 37° C with 5% CO_2 until the cells were 60-80% confluent.

12µg of DNA construct was mixed with 150µl of DMEM minus FBS, 115µl of Polyfect (Qiagen) added and left at RT for 10 mins. The cells were washed in PBS which was replaced with fresh DMEM + 10% FBS. 1ml of media was added to the DNA/media/Polyfect mixture and transferred to the flask. The cells were incubated for 24 hrs at 37° C + 5% CO₂, the media was then replaced with DMEM minus FBS and the cells incubated for a further 24 hrs after which the media and cells were both collected for analysis.

2.2.4.2 *In vitro* translation

Synthetic capped mRNA was tested using the TnT Reticulolysate *in vitro* translation kit (Promega) to check translational efficiency. This was carried out to the manufacturers recommended protocol. 1 µl of reaction was mixed with 4 µl of nfH₂O and 5 µl of 2 x loading buffer containing β-mecaptoethanol, and heated to 100°C for 5 minutes to denature the protein. The sample was then run on a precast 4-12% Bis-Tris SDS PAGE gel (Invitrogen) for 1 hour at 200V. The gel was then fixed in 40% Methanol/10% Acetic acid for 30 minutes, and gel dried at 80°C for 2 hours. The gel was then exposed overnight to Fujifilm SuperRX medical X-ray film at –80°C, and developed in the morning.

2.2.4.3 Western Blot

Samples were loaded onto precast 4-12% Bis-Tris SDS PAGE gels (Invitrogen) and run for 1 hour at 200V using the manufacturers running buffer. The gel was then equilibrated in transfer buffer for 20 minutes, and then blotted onto PVDF membrane for 30 minutes at 15V using a semi dry blotter (BioRad). The PVDF membrane had been washed in methanol and dH₂O, then equilibrated with transfer buffer prior to blotting. The blotted membrane was then blocked overnight in 5% Marvel in TBS at 4°C with gentle agitation. The following day the membrane was washed in TBST for 30 minutes repeated 3 times, then sealed in a bag containing the primary antibody in 2.5% Marvel/TBST (Primary antibody dilutions ranged from between 1/500 to 1/2000). The membrane was placed at 4°C for 3 hours then washed in TBST 3 times, again for 30 minutes each time, then washed overnight in TBST at 4°C. The membrane was then exposed to the secondary antibody in 2.5% Marvel/TBST at a dilution of 1/2000 at 4°C in a bag for 2-3 hours. After washing the membrane a

further 3 times 30 minutes in TBST the blot was developed using the ECL detection kit (Amersham) and exposed to Fujifilm SuperRX X-Ray film for between 30 seconds and 10 minutes depending on the strength of the signal.

2.2.5 Wholemount *in situ* hybridisation

2.2.5.1 Probe synthesis

Plasmid DNA was cut at a 5' restriction site so that the promoter at the 3' end will produce an antisense transcript. The DNA was then cleaned by Proteinase K treatment, Phenol:Choloroform extracted, and then ethanol precipitated. The DNA was then resuspended in nfH_2O at a concentration of $1\mu\text{g}/\mu\text{l}$. The following reaction was then set up on ice. All reagents were obtained from Promega unless otherwise stated.

2.0 μl DNA @ $1\mu\text{g}/\mu\text{l}$

2.0 μl 10mM ATP

2.0 μl 10mM CTP

2.0 μl 10mM GTP

1.3 μl 10mM UTP

0.7 μl 10mM Digoxigenin-11-UTP (Roche)

4.0 μl 5 x Reaction buffer

1.0 μl RNAsin

2.0 μl RNA polymerase (Sp6, T7, or T3)

made to 20 μl with nfH_2O .

The reaction mix was placed at 37°C for 3 hours. This was followed by the addition of 1 μl of RQ DNase1 and a further 15 minute incubation at 37°C to remove the DNA template. The probe was then ethanol precipitated, resuspended in 10 μl nfH_2O , and checked on a 1% TAE agarose gel. The probe was then used at a concentration of between 0.5 and 1.0 $\mu\text{g}/\text{ml}$.

2.2.5.2 *in situ* hybridisation.

Embryos fixed in MEMFA and stored at -20°C in 100% methanol were brought to RT, and rehydrated using the following method, (all steps are carried out with gentle agitation unless otherwise stated)

75% Ethanol: H_2O 10 minutes

50% Ethanol:H ₂ O	10 minutes
25% Ethanol:H ₂ O	10 minutes
PBSAT	5 minutes x 2

At this stage the embryos were either bleached in 6% H₂O₂, with Formamide and 2xSSC for 10 to 15 minutes on a light box, or treated with Proteinase K for 3 minutes. The two were not done together as this weakens the embryos, resulting in their disintegration. Following both procedures the embryos were washed in PBSAT for 3 x 5 minutes. The embryos were then placed in 5 ml of freshly made 0.1M Triethanolamine (pH 7.8 with Acetic acid) (Fisher), for 5 minutes. After a further 5 minutes in fresh Triethanolamine (5 ml), 12.5 µl of Acetic Anhydride (Sigma) was added. This was swirled occasionally for 5 minutes before a further 12.5 µl of Acetic Anhydride was added, and again swirled for 5 minutes. This step removes the positive electrostatic charges, which can affect the binding of the probe during hybridisation. The embryos were washed twice for 5 minutes in PBSAT, and then re-fixed in 10% Formalin for 20 minutes. After re-fixation the embryos were washed approximately 5 times for 5 minutes each in 5 ml of PBSAT, after the final wash 4 ml of PBSAT was removed and 250 µl of room temperature hybridisation buffer was added. The embryos were left to settle then the solution replaced with 1 ml of fresh hybridisation buffer. The embryos were then preincubated in this buffer at 4°C in a refrigerator for 2 to 3 days. After preincubation the embryos were removed from the refrigerator and warmed to room temperature, before being placed at 60°C for 10 minutes. The hybridisation buffer was replaced with fresh, prewarmed to 60°C and the embryos were left at this temperature for 4 to 6 hours. Before the addition, the probe was heated to 80°C for 3 minutes which denatures any secondary structure that the probe may have formed with it's self thus allowing better chance of hybridising to it's target. The probe was added to 1-2ml of fresh prewarmed 60°C hybridisation buffer in a final concentration of approximately 0.5 µg/ml, this was then added to the embryos, and left over night with gentle agitation in a 60°C incubator. The following morning the probe solution was removed and stored at -80°C for future use, (probes can be used a number of times before they stop detecting signal in the embryo). The following protocol was then used to wash the embryos, and remove any unbound probe to prevent background staining from trapped probe. All steps were carried out at 60°C with gentle agitation, and all solutions prewarmed to 60°C.

2 ml Hybridisation buffer	2 x 10 minutes
2 ml 2 X SSC + 0.1% Tween-20	3 x 30 minutes
2 ml 0.2 X SSC + 0.1% Tween-20	2 x 30 minutes

After these steps the embryos were washed twice in 2 ml of Maleic acid buffer (MAB) with gentle agitation for 15 minutes at room temperature. The embryos were then blocked with 2 ml of 2% Boehringer Mannheim blocking agent dissolved in MAB for 1 to 2 hours, and then preincubated for 1.5 to 2 hours in 2 ml of 2% block in MAB (as previously) containing 20% heat treated lamb serum (GibcoBRL) again at RT. The preincubation buffer was replaced with 2 ml of fresh containing a 1/2000 dilution of affinity purified sheep anti-digoxigenin antibody coupled to alkaline phosphatase. The embryos were then incubated over night with gentle agitation at 4°C. The next day the antibody solution was removed. The embryos were washed 3 times for 5 minutes in MAB, and then washed once an hour throughout the day again in MAB using 5 ml each time. Finally the embryos were washed in MAB over night at 4°C. The MAB was removed and embryos washed once for 5 minutes and once for 10 minutes with freshly made alkaline phosphatase buffer containing levamisole. This solution activates the alkaline phosphatase conjugated to the antibody. The levamisole inactivates endogenous alkaline phosphatase that is found in differentiated notochord. The buffer was replaced with BM purple alkaline phosphatase substrate (Roche), pre-warmed to RT, and left to develop at room temperature until the colour appears. The BM purple acts as a substrate for alkaline phosphatase and results in a blue/purple colour, the colouration corresponds to regions where the probe is localised, which in turn shows the location of the mRNA transcript of interest. Once the colour had developed the embryos were washed in PBSAT twice for 15 minutes, and then fixed in 10% Formalin for storage. Finally if a Proteinase K step had been included the embryos were bleached in 6% H₂O₂ on a light box after re-fixation. For particularly deep staining or to improve visualisation of a signal the embryos were cleared using 1,2,3,4-tetrahydronaphthalene (Sigma). After fixation in 10% Formalin, the embryos were placed in 100% Methanol for 5 minutes, Propan-2-ol for 5 minutes, and then placed in 1,2,3,4-tetrahydronaphthalene. These steps were completed in glass vials as 1,2,3,4-tetrahydronaphthalene corrodes plastic.

2.2.6 Wholemount antibody staining

Embryos fixed in MEMFA and stored in 100% Methanol were brought to RT, and rehydrated using the following scheme. All steps were carried out in small plastic tubes with gentle agitation at RT unless otherwise stated.

75% ethanol/TBST 10 minutes

50% ethanol/TBST 10 minutes

25% ethanol/TBST 10 minutes

TBST 5 minutes

TBST 5 minutes

(TBST: 50 ml 10 x TBS, 750 µl Triton X-100, made to 500 ml with dH₂O)

The embryos were then bleached on a white light box in 6% H₂O₂ in TBST for 1 to 2 hours, then washed in TBST twice for 5 minutes. The protein free ends were blocked with 1M glycine pH'd to 7.5 with ammonia for 30 minutes, then the embryos washed three times for 30 minutes each in TBSBT (TBS containing BSA at 2 mg/ml). The embryos were blocked in TBSBT containing 10% heat treated lamb serum for 2 hours, and incubated overnight at 4°C in fresh solution containing the primary antibody (dilutions used indicated in the text). The following day the embryos were washed 5 times for 1 hour in 4-5 ml TBSBT, and then for 1 hour in TBSBT containing 10% lamb serum. The secondary antibody, usually in a dilution of 1:1000, was added to 2 ml fresh TBSBT:10% lamb serum, and again incubated overnight at 4°C. The next day the embryos underwent further washing steps with 4-5 ml of TBSBT throughout the day each lasting an hour, and finally washed overnight in TBSBT at 4°C. The colourmetric reaction using 1 ml of BM purple alkaline phosphatase substrate (Roche) was carried out the next day after one 3 minute and one 10 minute wash in 2 ml of alkaline phosphatase buffer containing Levamisole. The reaction was left at room temperature until the colour developed, then the embryos were washed twice in 4-5 ml of TBST, and finally fixed in 10% Formalin.

2.2.7 Histology

2.2.7.1 Wax embedding of embryos

Embryos to be embedded and sectioned for histological analysis were fixed in Romesis fixative overnight, then placed in Borax-Carmine stain for a further night. In order to remove the excess Borax-Carmine leaving the nuclei stained red, the embryos

were placed in 70% ethanol containing 1% HCl for 4 to 6 hours. The embryos were then subjected to the following procedure to embed them in wax.

70% EtOH	30 minutes
90% EtOH:BuOH (1:1)	30 minutes
100% EtOH:BuOH (1:3)	30 minutes
100% BuOH	30 minutes

Fibrowax (BDH) melted at 60°C was added in a 1:1 ratio with the 100% BuOH, and the embryos left at 60°C for 30 minutes. The wax was replaced 3 to 4 times with fresh, each time incubating the embryos at 60°C for at least 30 minutes. Finally fresh wax added and the embryos incubated overnight at 60°C. The following day the embryos were placed in moulds and the wax allowed to set.

2.2.7.2 Sectioning

The wax blocks containing embedded embryos were trimmed and attached to wooden blocks. The embryos were then sectioned to thickness of 12 µm using a Microtome. The sectioned embryos were then floated onto poly-L-lysine subbed microscope slides and allowed to dry overnight on a hot plate at 45°C.

2.2.7.3 Counter staining

Sectioned slides were subjected to the following procedure. These steps remove the wax surrounding the embryo, and counter stain the cytoplasm of the cell, which results in a green colouration.

Histoclear (Fisher)	5 minute x 2 (Fresh Histoclear used for each wash)
100% EtOH	1 minute
90% EtOH	1 minute
70% EtOH	1 minute
50% EtOH	1 minute
dH ₂ O	1 minute
Picoro Blue/Black	1 minute
dH ₂ O	1 minute
70% EtOH	1 minute
90% EtOH	1 minute
100% EtOH	1 minute

Histoclear 1 minute x 2

A small drop of DPX histological mount (Fisher) was placed on the slide and a cover slip laid over the top. The slides were then left to dry at RT.

2.2.8 Electron microscopy

Embryos were fixed in 2.5% glutaraldehyde, 4% paraformaldehyde in 0.1M cacodylate buffer (pH 7.3), post fixed in 1% osmium tetroxide, EN-block stained in 2% uranyl buffer, dehydrated and embedded in Araldite resin. 70-90nm sections were collected on 300 mesh copper grids, stained in lead citrate, and viewed on a JOEL 1010 T.E.M.

2.2.9 Photography.

Whole embryos were photographed using a Nikon Coolpix 990 digital camera attached to a Leica Wild MZ5 dissecting microscope. All photographs were taken and stored in TIFF format with a resolution of 300dpi. Colour levels and contrast were adjusted using Adobe Photoshop 7.0. Autoradiographs and agarose gel photographs were scanned into Adobe Photoshop using either a UMAX or Canon flatbed scanner. Scanned gels were adjusted using colour level settings to improve the contrast of some images. All figures were assembled using Adobe Photoshop.

Chapter 3

Identification and characterisation of three members of the Lysyl oxidase family expressed in the notochord of developing *Xenopus laevis* embryos

3.1 Introduction

3.1.1 Lysyl oxidase

Lysyl oxidase (Lox)¹ is an extra-cellular copper dependent amine oxidase that catalyses the cross-linking of collagen and elastin fibres in the extracellular matrix (ECM). By deaminating the ϵ -amino group of peptidyl-lysines in collagen and elastin, Lox allows the spontaneous condensation of these fibres into mature fibrils (Csiszar, 2001; Kagan, 1986). First translated as a 46kDa prepro-enzyme, Lox undergoes N-glycosylation and is secreted as a 50kDa inactive pro-enzyme. A 14kDa N-terminal pro-peptide is then removed by BMP-1 to yield an enzymatically active form of Lox (Panchenko et al., 1996; Uzel et al., 2001).

The primary function of Lox is thought to be the cross-linking of collagen and elastin fibres, as demonstrated by mouse null mutants that had reduced collagen and elastin cross-links (Hornstra et al., 2003; Maki et al., 2002). However, a number of recent studies suggest other functions for the enzyme, including a number of intracellular roles. A fully processed form of Lox is able to translocate from the ECM to nuclei of aortic smooth-muscle cells (Nellaiappan et al., 2000) and Lox activity has been detected within the nuclei of fibrogenic cells (Li et al., 1997). Lox can also activate the human *collagen type-III* promoter (Giampuzzi et al., 2000) and reduce the binding

¹ In this chapter I will use the abbreviation Lox to describe lysyl oxidase itself and the abbreviation LOX to describe both the conserved catalytic domain and the family of lysyl oxidase related enzymes.

and activity of NF-kappa- β (Jeay et al., 2003). It has also been reported that Lox has anti-oncogenic activity on ras transformed cells (Giampuzzi et al., 2003), although a recent study indicates that the pro-peptide, yielded by BMP-1 cleavage, is responsible for this action (Palamakumbura et al., 2004). Whether Lox catalytic activity or the action of its pro-peptide is responsible for other functions described remains to be seen.

3.1.2 Lox-like enzymes

In mammals, four genes closely related to *lox* have been identified. To date the family consists of *lox* (Hamalainen et al., 1991), *lysyl oxidase-like 1 (loxl-1)* (Kenyon et al., 1993), *lysyl oxidase-like 2 (loxl-2)* (Saito et al., 1997), *lysyl oxidase-like 3 (loxl-3)* (Jourdan-Le Saux et al., 1999; Maki and Kivirikko, 2001) and *lysyl oxidase-like 4 (loxl-4)* (Asuncion et al., 2001; Maki et al., 2001). Each member of the *lox* family contains a highly conserved C-terminal domain consisting of a copper-binding motif, residues for lysyl-tyrosylquinone (LTQ) co-factor linkage, and a cytokine receptor-like domain. The formation of an LTQ co-factor is required for Lox activity and forms when the ϵ -amino group of a lysine residue (Lys³¹⁴ in humans) interacts with the modified side chains of a Tyrosine (Tyr³⁴⁹) (Wang et al., 1996). The Cu²⁺ ion acts as a second co-factor, its removal decreases the activity of Lox and it is thought to play a role in stabilising the LTQ (Gacheru et al., 1990). However, the LTQ domain is still able to form and Lox retains some activity in the absence of Cu²⁺ (Tang and Klinman, 2001).

In contrast to the catalytic domain, the N-terminal region of the Lox-like proteins differs considerably in relation to each other. Loxl-1 contains a 100 amino acid

proline rich domain, while Loxl-2, -3, and -4 contain four repeated scavenger-receptor cysteine-rich (SRCR) motifs (Csiszar, 2001) (Figure 3.1). The highly variable N-terminal domains may be responsible for any differences in function exhibited by the different family members

Comparatively little is known about the function of the Lox-like enzymes, but the amino acid sequence of the C-terminal catalytic domain is highly conserved suggesting that they all share the same enzymatic activity. Indeed, Loxl-1 and Loxl-4 have been shown to have similar amine oxidase activity to Lox *in vitro* using collagen as a substrate (Borel et al., 2001; Kim et al., 2003). However, Loxl-1 appears to have a different function from Lox *in vivo*. Murine Loxl-1 localises to areas of elastin fibre formation and associates with the elastin associated component fibulin-5 (Liu et al., 2004). Mice that are homozygous for a null mutation of the *Loxl-1* gene display abnormalities in elastin rich structures, such as the uterine tract, but collagen fibres remain unaffected (Liu et al., 2004).

Lox genes are not confined to vertebrates with two members of the Lox-like family having also been identified in *Drosophila* (*DmLoxl-1* & *DmLoxl-2*) (Molnar et al., 2005). *DmLoxl-1* is expressed during embryonic development but functional studies have not yet given a conclusive role for this family member. Lox activity has also been

Figure 3.1 A schematic representation of the Lox family of proteins

All Lox enzymes contain a C-terminal catalytic domain (blue box) and a more variable N-terminal region. The N-terminal region of Lox itself does not contain many characteristic features. Loxl-1, however, contains a proline rich region (red box), while Loxl-2, 3, and 4 contain four Scavenger Receptor Cysteine Rich (SRCR) repeats and are significantly longer at approximately 774 amino acids.

detected in sea urchin embryos where it is thought to have a role in regulating cell movements during gastrulation (Gosline et al., 1974; Wessel and McClay, 1987). The

Figure 3.1 The sea urchin *Lox* activity have yet to be described.

3.1.3 Inhibitors of *Lox* function

There are a number of chemical inhibitors of *Lox* activity, including the organophosphate Malathion (Shawyer and Chambers, 1993), Benzoyl Hydrazine

(Riggin and Shultz, 1986) and the nitriles, β -Aminopropionitrile (β -APN) and Samacarbazide (Schultz et al.

effects using the frog embryo ventral fin assay (FETA) *Lox* inhibition occurs as a result of improper

β -APN is a specific irreversible inhibitor of the *Lox* (Schultz et al., 1997). When

β -APN is applied to sea urchin embryos it causes the arrest of gastrulation (Butler et al., 1987; Wessel and McClay, 1987) and endoderm formation appears to be disrupted

(Goslin et al., 1997). In frog embryos the major defects caused by β -APN include disruption of the notochord and its surrounding sheath, both collagen rich structures

(Dawson et al., 2000; Moutier et al., 1997). These studies, however, have only investigated the toxicity of β -APN using FETA and do not address any possible

patterning defects caused by *Lox* inhibition.

3.1.4 Summary

This chapter presents data identifying three *lox* genes that are expressed during *X. laevis* embryogenesis. They are most closely related to mammalian *lox*, *loxl-1* and

loxl-3. Transcripts for all three genes are localized to the notochord from early

detected in sea urchin embryos where it is thought to have a role in regulating cell movements during gastrulation (Butler et al., 1987; Wessel and McClay, 1987). The genes responsible for sea urchin Lox activity have yet to be described.

3.1.3 Inhibitors of Lox function

There are a number of chemical inhibitors of Lox activity, including the organophosphate Malathion (Snawder and Chambers, 1993), Benzoyl Hydrazine (Riggin and Shultz, 1986) and the nitriles, β -Aminopropionitrile (β -APN) and Semicarbazide (Schultz et al., 1985). All were identified as causing osteolathrytic effects using the frog embryo teratogenesis assay (FETAX). Osteolathyrisms occur as a result of improper collagen formation, one cause being defective Lox action. β -APN is a specific, irreversible inhibitor of the Lox catalytic domain (Tang et al., 1983) and is the toxic compound found in the sweet pea *Lathyrus odoratus*. When β -APN is applied to sea urchin embryos it causes the arrest of gastrulation (Butler et al., 1987; Wessel and McClay, 1987) and endoderm formation appears to be disrupted (Godin et al., 1997). In frog embryos the major defects caused by β -APN include disruption to the notochord and its surrounding sheath, both collagen rich structures (Dawson et al., 2000; Mentzer et al., 1999). These studies, however, have only investigated the toxicity of β -APN using FETAX and do not address any possible patterning defects caused by Lox inhibition.

3.1.4 Summary

This chapter presents data identifying three *lox* genes that are expressed during *X. laevis* embryogenesis. They are most closely related to mammalian *lox*, *lox1-1* and *lox1-3*. Transcripts for all three genes are localised to the notochord from early

neurula stages, but expression of *Xlox* subsequently expands into the somites and head. To elucidate the role of Lox like enzymes during development, embryos were treated with the specific chemical inhibitor β -APN. Treated embryos appeared to develop normally until late tailbud stages, when kinks in the notochord appeared and the somites slipped towards the tail. Blisters subsequently appeared in the epidermis and the gut failed to coil and differentiate. EM sections showed that β -APN disrupted the formation of collagen fibres in the notochord sheath, indicating that disturbing the development of connective tissues causes these defects.

3.2 Results

3.2.1 Identification of three *lox* genes during *X. laevis* development

Searching the NCBI EST databases for potential substrates of the Tld metalloproteases, three *X. laevis* cDNAs encoding different members of the *lysyl-oxidase* family were identified. Sequencing and subsequent analysis of the clones revealed two cDNAs most closely related to *lysyl-oxidase* (*Xlox*) and *lysyl oxidase-like 1* (*Xloxl-1*), and a smaller, partial length, clone most closely related to *lysyl oxidase-like 3* (*Xloxl-3*). *Xlox* encodes a 429 amino acid protein (XLox) that is 63% identical to both mouse and human Lox (mLox and hLox), and 66% identical to chick Lox (cLox). XLox contains a putative cleavage site for BMP-1 between amino acids 175 and 180 (Figure 3.2A clear box) and a C-terminal lysyl-oxidase (LOX) catalytic domain (amino acids 297-426, Figure 3.2A shaded box). The LOX domain is highly conserved between species and a comparison of this domain alone shows 89%, 90% and 93% identity to the mLox, hLox, and cLox domains respectively. *Xloxl-1* encodes a 528 amino acid protein that is 52% and 54% identical to mLoxl-1 and hLoxl-1 respectively. In addition to the highly conserved LOX domain with copper binding motif (Figure 3.2B blue box), XLoxl-1 contains a proline rich region (amino acids 193-294, clear box) that appears to account for the 100 amino acid difference between XLox and XLoxl-1. In contrast to mLoxl-1 and hLoxl-1, XLoxl-1 does not have a conserved BMP-1 cleavage site (Figure 3.2B). The third clone encodes 299 amino acids that aligned most closely to the C-terminal region of mouse and human Loxl-3 (mLoxl-3, hLoxl-3), between amino acids 450 and 755. We have designated this clone *Xloxl-3*. A characteristic feature of Loxl-3 from all species so far studied is the presence of four Scavenger Receptor Cysteine Rich (SRCR) domains (see Figure 3.1).

Figure 3.2 Deduced protein sequences and alignments for the *Xenopus* Lox family

The deduced amino acid sequences for Xlox, XLoxl-1, and XLoxl-3 were aligned to their human, mouse, and chick (where available) orthologues. In all aligned sequences the highly conserved catalytic domain (shaded box), copper binding motif WxWHxCHxHYHSMD (blue box), and the lysine and tyrosine residues required for LTQ co-factor linkage (red letters) are indicated.

A. XLox is a 429 amino acid protein that shares 63% identity with mouse (mLox) and human (hLox) Lox, and 66% identity with chick Lox (cLox). A putative cleavage site for BMP-1 (clear box) is conserved in all four proteins.

B. XLoxl-1 is a 528 amino acid protein with 52% and 54% identity to mLoxl-1 and hLoxl-1 respectively. The proline rich domain is indicated (clear box).

C. Partial amino acid sequence (299 amino acids) for the C-terminus of XLoxl-3. This sequence shares 83% and 81% identity to mLoxl-3 and hLoxl-3 respectively. The first 72 amino acids are homologous to the fourth SRCR domain of the mammalian proteins.

Figure 3.2A

XLox	1	MGCAPGRVFLQLVLVYLHCLLLGDCQQQRSPVGPSSMWRQLQWQSNGRV	50
MmLox	1	.RF.WAVLL.GP.Q.CPLLRCAPQTPREPP--AAPGA...TI..EN..Q.	48
HsLox	1	.RF.WAVLL.GP.Q.CPLLRCAPQTPREPP--AAPGA...TI..EN..Q.	48
GgLox	1	.R...PGLL.A..HACIFWSG.WPAGC.----S.PAA.....N..Q.	46
XLox	51	YSLLSHGSEYQPARRE-----QEGSAPHSPLLLLSSRHNNNSVSQ	90
MmLox	49	F....L.AQ...Q...DPSATARRPDGDAA.Q.RT.I...RDN-----	91
HsLox	49	F....L.AQ...Q...DPSATARRPDGDAA.Q.RT.I...RDN-----	91
GgLox	47Q.AQ...P...Q-----AE.A...V...RGNGS----	80
XLox	91	GTGAHRAQPPRSVTPSAEGQQQQPASTNPAAREDAATSRLLWFQATRQRN	140
MmLox	92	-----TAST.AR...PS.VAAGRPRP-----AA.....GFSP-	125
HsLox	92	-----TAST.AR...PS.VAAGRPRP-----AA.....GFSP-	125
GgLox	81	---VP.AAAAAARPQP.P.P.AQPQPR.RSSRRQPLGR.....GYRAP	127
XLox	141	AG-DAAGRQRSAGAAQERRQ-----GAQASHTRVNGTDRVEGMOGDDP	182
MmLox	126	S.-ARD.AS....NRTASP.P-----P.L.NL.PP--SHIDR.V....	165
HsLox	126	S.-ARD.AS....NRTASP.P-----P.L.NL.PP--SHIDR.V....	165
GgLox	128	S.SARP.PR..PRGRRSR..ERAERRRAAAP.GL.P---GREDV.V....	174
XLox	183	YNPKYSEDNPYYNYDYDTERSRTNERQRPYGTRYFQNGLPDLVPDPYY	232
MmLox	166P.PGS.N.....G...Y.....	215
HsLox	166P.PGS.N.....G...Y.....	215
GgLox	175Q..RY.....G...Y.....	224
XLox	233	IQASIYVQKMSMYNLRCAAEENCLASSAYSSDVRYDQRVLLRFPQRVKN	282
MmLox	216RA.....H.....	265
HsLox	216	...T.....RA.....H.....	265
GgLox	225	...T.....RA.....N.....	274
XLox	283	QGTADFLPSRPRYTWEWHSCHQHYHSMDEFSHYDILLDASSHRRVAEGHKA	332
MmLox	266	...S.....S.....NTQ.....	315
HsLox	266	...S.....S.....NTQ.....	315
GgLox	275	...S.....S.....	324
XLox	333	SFCLEDTSDDYGYRRFACTAHTQGLSPGCYDTYNADIDCQWIDITDVKP	382
MmLox	316H...G.....A.....Q.	365
HsLox	316H.....A.....Q.	365
GgLox	325Y..Y.C.....	374
XLox	383	GNYLKVSVPNSYQVPESDYSNNVVRCDLRYTGHYVYTSGCTTISPY*	429
MmLox	366L.....T.....HA.A.....*	412
HsLox	366L.....T.....HA.A.....*	411
GgLox	375L.....I....I....HA.A.....*	420

Figure 3.2B

Xloxl-1	1	MAFPG---QL-----CLWFWALSVAL--GQEGDS--WGQMIQWEDSGR	36
HsLoxl-1	1	..LARGSR..GALVWGA..CVLVHGQQAQP..GS.PAR.R.L...NN.Q	50
MmLoxl-1	1	..LA.AGS..RTLWWSA..CVLVHGQQAQP..GS.PGR.R.L...NN.Q	50
Xloxl-1	37	RYRLNLSGSEYQAAGGGTG--GSRVLLDGIRSGGS-----DLRRRQAPSS	79
HsLoxl-1	51	V.S.....VP..PQRSESS.....A.APQAQRRSHGSP.....L	100
MmLoxl-1	51	V.S.....VP..PQR.ETS.....A.APQTSQRRSQGGP.....L	100
Xloxl-1	80	P---RTSSQTVRGNTRHPFGFGQVPDNWRSVSESASTGRFVPSAGAG---	123
HsLoxl-1	101	.LPG.VG.D....QA.....E.AVGD...MALARTSVSQQ-	149
MmLoxl-1	101	.LPG.VG.D....Q.....E.AVGD...MARARTSVSQQR	150
Xloxl-1	124	--GGGGRIRQSSSQASVG-----TQFFPRPQPPFVPQTDLNP---QGYD	162
HsLoxl-1	150	RH..SASSVSA.AF..TYRQQPSYP.Q..Y..A...S.YENYDPASRT..	199
MmLoxl-1	151	HG.SASSSVSA.AF..TTYRQPSYP.Q..Y..A...N.YENYDPASRT.E	200
Xloxl-1	163	EGYGYQRAGGGGGSYYAQSWAGGGYEDAWEEPSPPFAQPPY--FGIPPNL	210
HsLoxl-1	200	Q.FV.Y.PA...VGAG.AAV.SA.VIYPYQPRARYEYGGG--EEL.EYP	247
MmLoxl-1	201	Q..V.Y.GA...MGAG.AAV.SA.VIYPFQPRARYEDYGGGGGEEQ.EYP	250
Xloxl-1	211	P-----PQQAQANPVV-----PQDGLDRRFAHSLFRGEDPPV	243
HsLoxl-1	248	.QGFYPA.ERPYPVP.PP-----P.....YS...YSEGT.GF	286
MmLoxl-1	251	AQGFYPA.ERPYPVQ.QPQPQPQPQPQ.S.....YS...YNEG.T.GF	300
Xloxl-1	244	VASEPVR-----PGYGGGGDTGFPPFGGARETDG-GYYGSTRP---	280
HsLoxl-1	287	EQAY.DP-----GPEAAQAHG.D.RL.WYPPYAN-PPPEAYG.PRA	326
MmLoxl-1	301	EQAY.DPSTDVSQAPAGAG.TY.GA.D.RL.WYPPYAANVPPEAYV.PRA	350
Xloxl-1	281	-----EPVLPPARVPSAPASGGQAQVEQQGRTSVGNVFRGNQNGR	320
HsLoxl-1	327	L-----PYL.V.SSDT.PP..ERNGA....L...S.Y.P....	367
MmLoxl-1	351	VEPQPPFRVL..PYL.V.SSD..SQ..ERNGA....L...S.Y.P....	400
Xloxl-1	321	GLPDLVPDPSYVQAATYIQRALYSLRCAAEENCLSSSAYAAEATDYDVR	370
HsLoxl-1	368N....S..V.....K..A.T.....	417
MmLoxl-1	401N....S..V.....K..A.T...P.....L.	450
Xloxl-1	371	VLLRFPQRVKNQGTADFLPTRPQSWEWHSCHQHYHSMDEFESHYDILLDAT	420
HsLoxl-1	418N....HT.....A	467
MmLoxl-1	451N....HT.....S	500
Xloxl-1	421	TGRVAEGHKASFCLDTCDFGNLKRYACTSHTQGLSPGCYDTYNADID	470
HsLoxl-1	468	..K.....S.....	517
MmLoxl-1	501	..K.....S.....	550
Xloxl-1	471	CQWIDITEVKPGNYILKVVVNPKYKVLESDFTNVVRNIHYTGRIYASAT	520
HsLoxl-1	518D.Q.....I.....V...	567
MmLoxl-1	551D.Q.....H.....I.....V.T.	600
Xloxl-1	521	NCRI TQF*	528
HsLoxl-1	568	..K.V.S*	575
MmLoxl-1	601	..K.V.S*	608

Figure 3.2C

Xlox1-3	1	WIPRAAGIRHE----GYSLHAVTETWYWDASNTTEMVLSGVRCRTGREMSL	46
HsLox1-3	450	.GTLEAMVACRQLGL..AN.GLQ.....SG.I..V.M.....T.L..	500
MmLox1-3	450	.GTLEAMVACRQLGL..AN.GLQ.....SG.V..V.M.....S.L..	500
Xlox1-3	47	EQCSHH-SSVSCKNTGTRHAAGVICSETASDLVLHSSLVQETAYIEDRPL	95
HsLox1-3	501	D.CA..GTHIT..R....FT.....L...A.....	550
MmLox1-3	501	N.CA..S.HIT..K....FT.....L...A.....	550
Xlox1-3	96	HMLYCAAEENCLSSSARHANWPYGHRRLLRFSSQIHNNGRADFRPKAGR	145
HsLox1-3	551A....S.....L.....	600
MmLox1-3	551A....S.....L.....	600
Xlox1-3	146	SWVWHECHGHYHSMDIFTHYDMLTPNGTKVAEGHKASFCEDESECQELVS	195
HsLox1-3	601I.....T...D..	650
MmLox1-3	601I.....T...D..	650
Xlox1-3	196	KRYECANFGEQGITVGCWDLWRHDIDCQWIDITDVKPGNYILQVVINPNF	245
HsLox1-3	651	700
MmLox1-3	651	700
Xlox1-3	246	EVAESDFTNNAMKCNCKYDGHRIWIHNCHLGDAFSEEANKRFEHYRGELN	295
HsLox1-3	701I.....R...R.P.QTS	750
MmLox1-3	701I.....R...R.P.QTS	750
Xlox1-3	296	NQIL* 299	
HsLox1-3	751	...I* 755	
MmLox1-3	751	...V* 755	

The clone described here contains the last 72 amino acids of a SRCR domain located N-terminally to the LOX domain (Figure 3.2c, amino acids 1-72). This SRCR domain is most likely to be the fourth SRCR domain; it is 56% and 59% identical to the fourth SRCR domain from HsLoxl-3 and MmLoxl-3 respectively, while only displaying 35% and 39% identity to the last 72 amino acids of the fourth SRCR domains from HsLoxl-2 and HsLoxl-4. The LOX domain of each protein (highlighted grey) contains a putative copper binding motif (WxWHxCHxHYHSMD) as identified by Prosite (Sigrist et al., 2002) (<http://www.expasy.org/Prosite>) (Figure 3.2A, B, C highlighted blue), a motif found in all lysyl oxidases, and conserved lysine and tyrosine residues for LTQ cross linkage (Xlox = Lys³³¹ & Tyr³⁶⁶; Xlox1-1 = Lys⁴²⁴ & Tyr⁴⁶⁵; Xlox1-3 = Lys¹⁷⁴ & Tyr²¹⁶; Figure 3.2A, B, C indicated in red). The full nucleotide sequence, highlighting the open reading frame of each clone, can be found in Appendix A, B and C.

3.2.2 *Xlox*, *Xlox1-1* and *Xlox1-3* are expressed in the notochord during *X. laevis* embryogenesis

The expression pattern of each clone was determined by both RT-PCR and whole mount *in situ* hybridisation. Expression of *Xlox* is first detected by RT-PCR at stage 13/14, and increases as development proceeds through stage 24/25 and into stage 35/36 (Figure 3.3A). *Xlox1-1* is detected by RT-PCR at stage 5/6 indicating the presence of maternal transcripts and is found at all the subsequent stages tested, with an increase in transcript levels during tailbud stages (Figure 3.3B). Transcripts for *Xlox1-3* are also first detected at stage 13/14 and at all subsequent stages tested with a slight increase as development

Figure 3.3 Temporal expression of *xlox*, *xloxl-1*, and *xloxl-3* as determined by RT-PCR

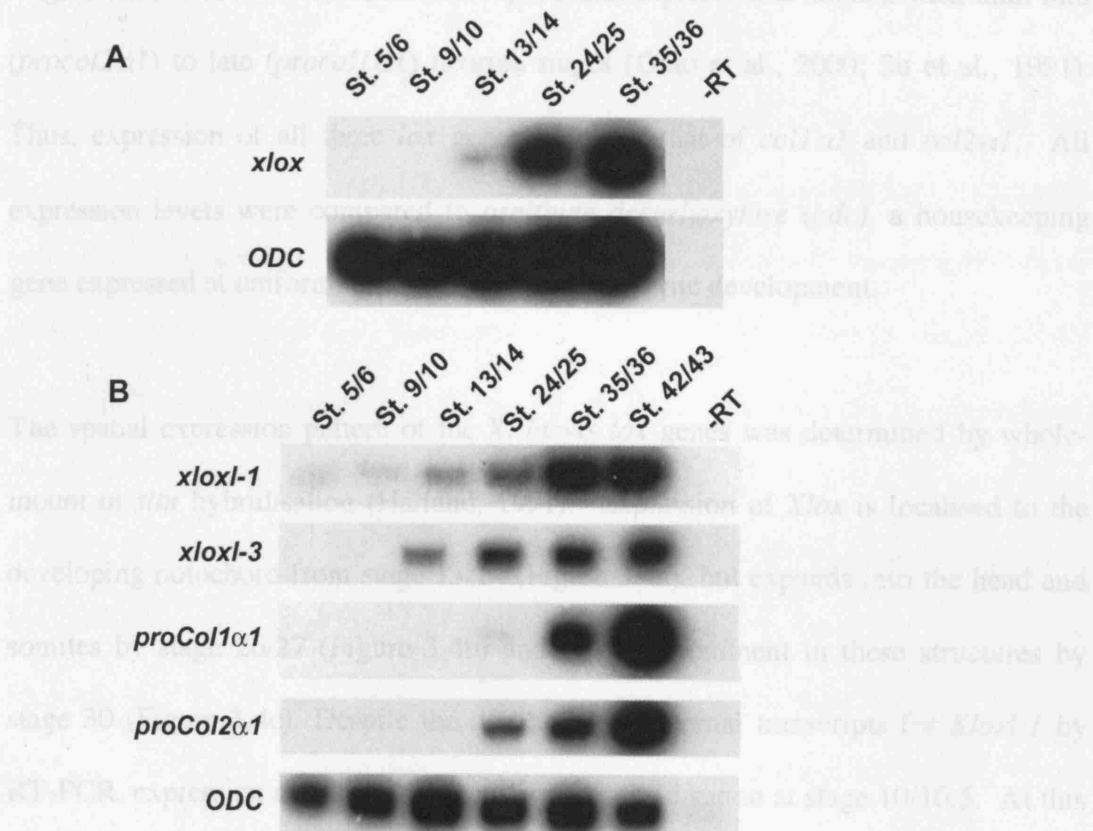
A. Expression of *Xlox* is first detected at stage 13/14 by RT-PCR and expression levels increase as development proceeds. **B.** Expression of *Xloxl-1* is first detected as maternal transcripts at stage 5/6, and increases as development proceeds. *Xloxl-3* is not detected maternally but is first expressed at stage 13/14, and again expression levels increase as development proceeds. Transcripts for *proColl α 1* and *proCol2 α 1*, which encode substrates for the Lox family, are not detected until stage 24/25. The uniformly expressed *ornithine decarboxylase (ODC)* gene was used as a loading control and total RNA from stage 24/25 embryos was used as a negative control. This result is representative result of at least three independent experiments.

proceeds. For comparison, we determined the expression of *procoll1 α* and *procoll2 α*

Figure 3.3

Both *Col1 α* for collagen synthesis that was expressed during early *X. laevis* development (Owen et al., 2000; Su et al., 1991) and are known substrates for Lox (Owen, 2001; Kagan, 1990). Transcripts for both genes are not detected until stage 24/25, consistent with a variety of reports that expression is not activated until mid (pmx/cr1) to late (procr) gastrula (Owen et al., 2000; Su et al., 1991).

Thus, expression of all three *Lox* genes, *Loxl-1*, *Loxl-2* and *Loxl-3*. All expression levels were normalized to *ODC* (Owen et al., 2000) a housekeeping gene expressed at similar levels throughout development.



Stage transcripts are localized to the mesenchyme of the dorsal marginal zone, i.e. the presumptive notochord (Figure 3.4d). For the remainder of development transcripts are detected only in the notochord (Figure 3.4e-h). Expression of *Loxl-3* is first detected at stage 13/14 and is also localized to the notochord at this and subsequent stages (Figure 3.4i-k). Transcripts for *Loxl-2* appear to be expressed more anteriorly, into the prechordal plate, then transcripts for *Loxl-1*.

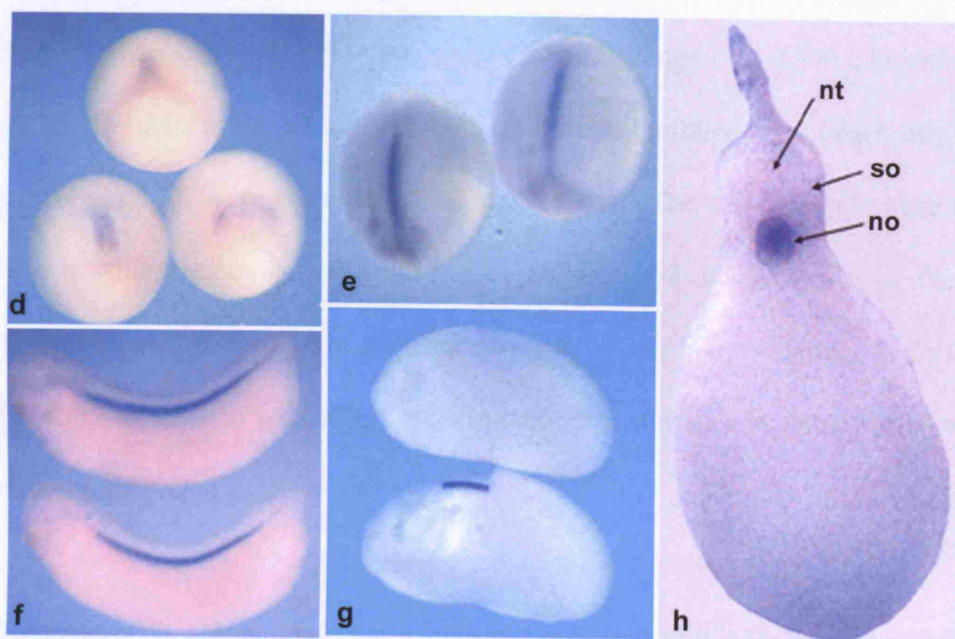
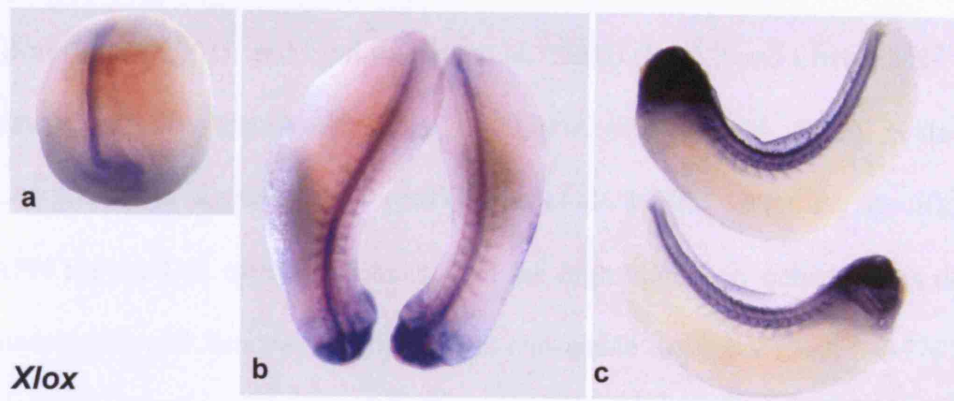
proceeds. For comparison we determined the expression of *procoll 1 α* and *procol2 α* (Figure 3.3B). Both code for fibrillar collagens that are expressed during early *X. laevis* development (Goto et al., 2000; Su et al., 1991) and are known substrates for Lox (Csiszar, 2001; Kagan, 1986). Transcripts for both genes are not detected until stage 24/25, consistent with previous reports that expression is not activated until mid (*procol2 α*) to late (*procoll 1 α*) neurula stages (Goto et al., 2000; Su et al., 1991). Thus, expression of all three *lox* genes precedes that of *coll 1 α* and *col2 α* . All expression levels were compared to *ornithine decarboxylase (odc)*, a housekeeping gene expressed at uniform levels throughout embryonic development.

The spatial expression pattern of the *X. laevis lox* genes was determined by whole-mount *in situ* hybridisation (Harland, 1991). Expression of *Xlox* is localised to the developing notochord from stage 13/14 (Figure 3.4a), but expands into the head and somites by stage 26/27 (Figure 3.4b) and is very prominent in these structures by stage 30 (Figure 3.4c). Despite the detection of maternal transcripts for *Xlox1-1* by RT-PCR, expression is first detected by *in situ* hybridisation at stage 10/10.5. At this stage transcripts are localised to the mesoderm of the dorsal marginal zone, i.e. the presumptive notochord (Figure 3.4d). For the remainder of development transcripts are detected only in the notochord (Figure 3.4e-h). Expression of *Xlox1-3* is first detected at stage 13/14 and is also localised to the notochord at this and subsequent stages (Figure 3.4i-k). Transcripts for *Xlox1-3* appear to be expressed more anteriorly, into the prechordal plate, than transcripts for *Xlox1-1*.

Figure 3.4 Expression of *Xlox*, *Xloxl-1*, and *Xloxl-3* determined by wholemount *in-situ* hybridisation

a-c. Transcripts for *Xlox* are first detected in the notochord at stage 14 (a) and continue to be detected in this structure at stage 25 (b) and stage 30 (c). Expression spreads into the somites and head at these later stages, being quite prominent in these structures at stage 30. **d-h.** Transcripts for *Xloxl-1* are first detected at stage 10/10.5 in a small band in the DMZ (d), in the region of the presumptive notochord. Expression is exclusively localised to the notochord at stages 14 (e), 25 (g), and stage 30 (f). A cross section through a stage 30 embryo (h) shows that expression is localised to the notochord (no), and is completely absent from the neural tube (nt) and somites (so). **i-k.** Transcripts for *Xloxl-3* are first detected at stage 14, once again in the notochord (i). *Xloxl-3* continues to be expressed in the notochord stage 25 (j) and 28 (k).

Figure 3.4



3.2.3 Inhibition of LOX function by β -APN treatment results in defects consistent with collagen malformation and does not affect A-P axis formation.

β -Aminopropionitrile (β -APN) is a specific inhibitor of the LOX catalytic domain, irreversibly blocking the activity of mammalian Lox (Tang et al., 1983), Loxl-1 (Borel et al., 2001), and Loxl-4 (Kim et al., 2003) (Loxl-2 and Loxl-3 have yet to be tested), as well as *Drosophila* Loxl-1 and Loxl-2 (Molnar et al., 2005). It also inhibits Lox activity in sea urchin embryos (Butler et al., 1987). Given the specificity of β -APN for the Lox catalytic domain, and the high homology between this domain in mammalian and *Xenopus* enzymes, it is reasonable to suppose that β -APN will also inhibit *Xenopus* lysyl oxidases. Consequently, *X. laevis* embryos were incubated from stage 6 (32-cells) through to stage 45 with a range of β -APN concentrations (1 μ M to 10 mM) in 5% NAM, replacing with fresh inhibitor media each day. Sibling control embryos were cultured in 5% NAM alone. The same defects were observed over a wide range of concentrations (10 μ M to 1 mM), while at the highest concentration (10 mM) embryos were distinctly sickly in appearance. Only 0.25 mM β -APN was used in subsequent experiments, the lowest concentration at which all embryos exhibited defects.

No obvious phenotypic defect was detected in 0.25 mM β -APN treated embryos until stage 41 (Figure 3.5F) when a number of kinks along the length of the notochord became apparent. The somites also appeared disrupted and there was a shortening in the length of the embryo. By stage 45 more severe abnormalities were visible, including severe shortening of the body and considerably more kinking of the notochord (compare Figure 3.5G & H). When analysed histologically, embryos treated with β -APN had formed fluid

Figure 3.5 Treatment of *Xenopus* embryos with 0.25mM β -APN

Embryos were cultured in either 5% NAM (A, C, E, G, I), or 5% NAM containing 0.25mM β -APN (B, D, F, H, J). There was no obvious phenotypic difference between untreated control embryos (A, C) and those treated with β -APN (B, D) at both stage 27/28 (A, B) and stage 33/34 (C, D). At stage 41, β -APN treated embryos (F) had developed a number of kinks in the notochord, and their A-P axis appeared shorter when compared to untreated controls (E). This shortening and kinking of the notochord was more apparent at stage 45 (compare G, H). Sections through the abdomen at stage 45 show that there is blistering of the skin and disruption to gut formation and coiling in β -APN treated embryos (Compare I & J). The result shown is representative of the phenotype observed in 100% of embryos in all experiments conducted (n= >100).

Figure 3.5

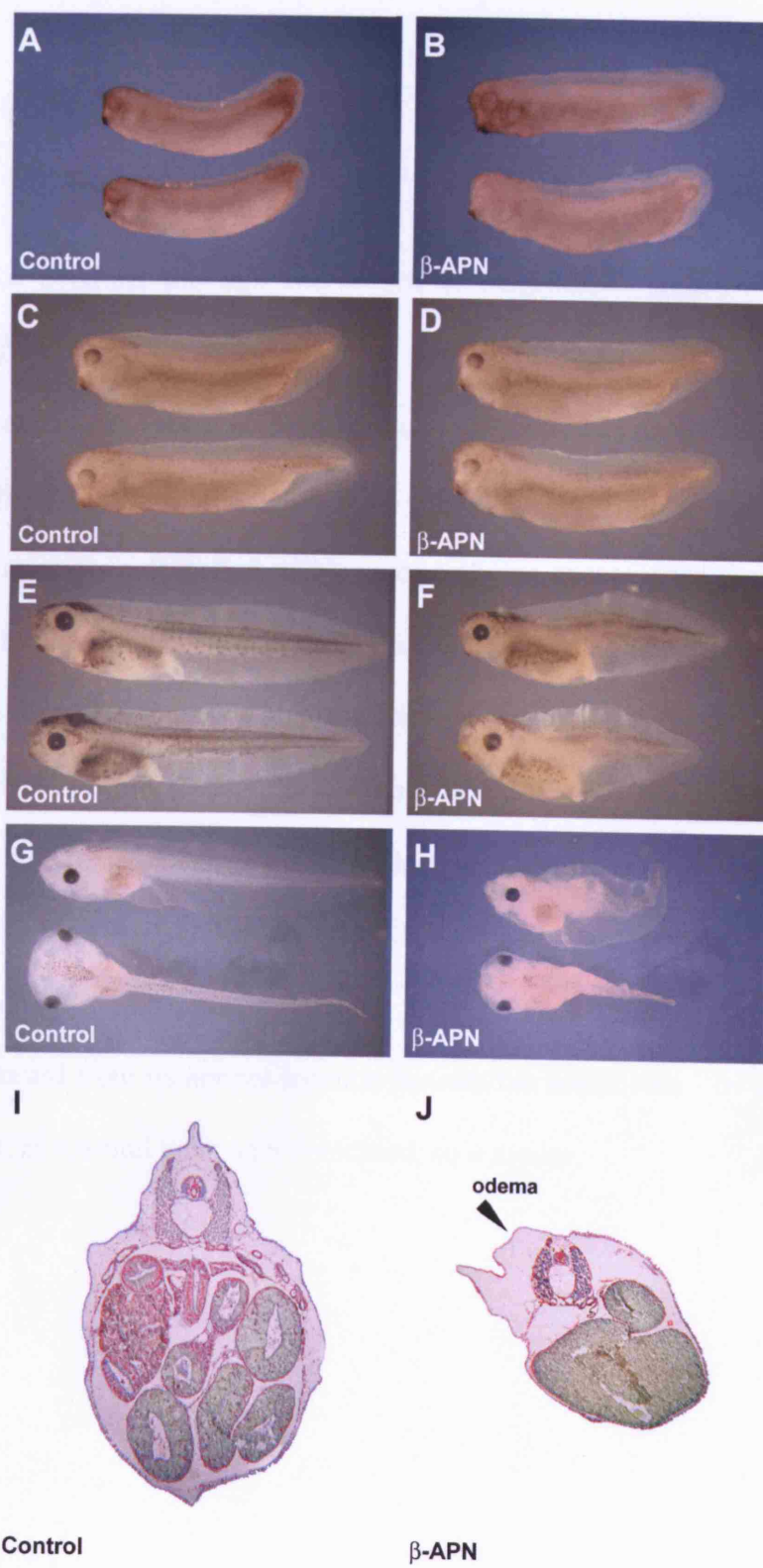
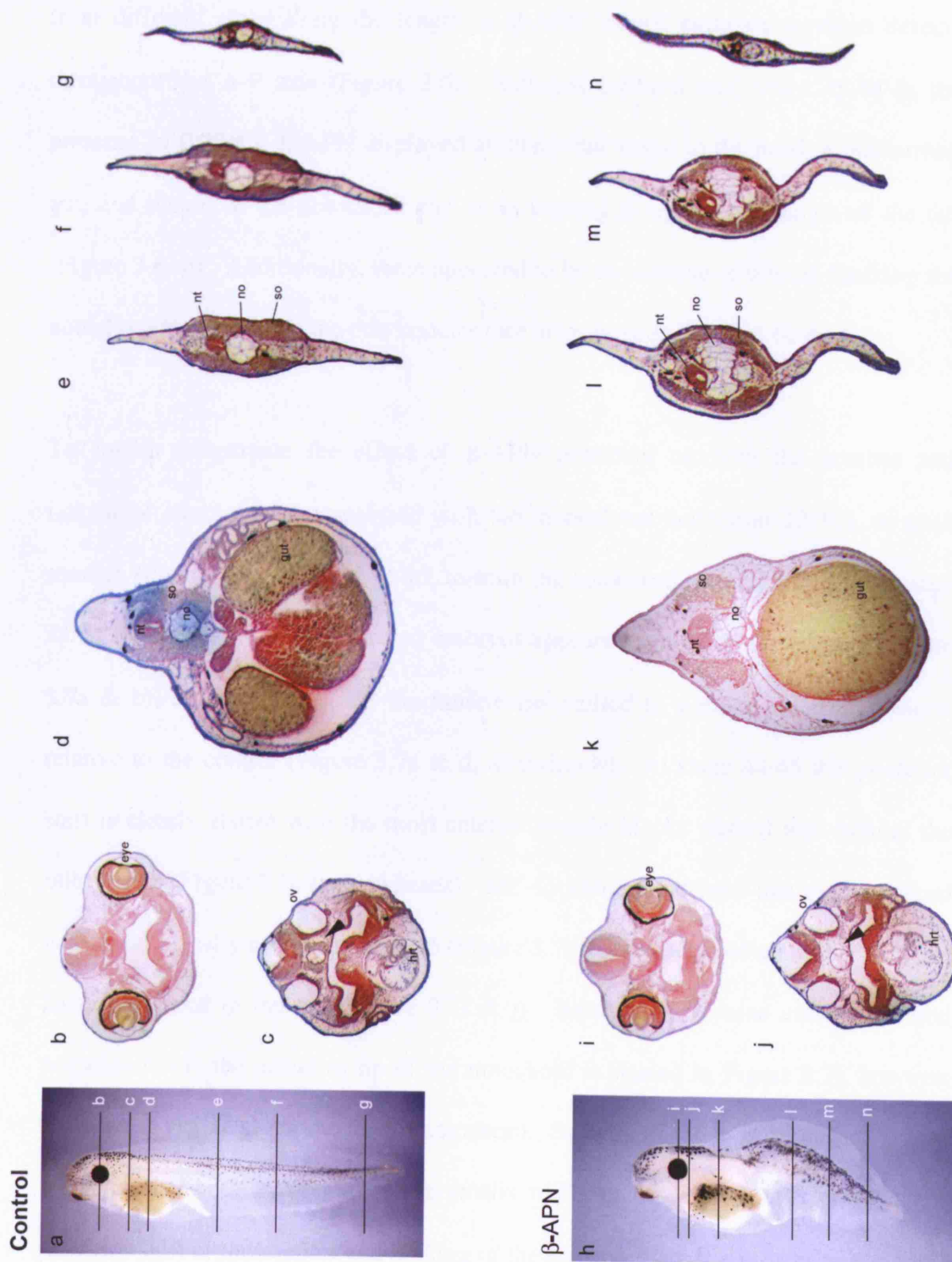


Figure 3.6 Sections through the length of stage 43/44 embryos treated with 0.25mM β -APN

Sagittal sections were taken at points throughout the length of both control (a-g) and 0.25mM β -APN treated (h-n) stage 43/44 embryos. A section through the eyes (b & i) reveals no major defects in β -APN treated embryos, except in the shape of the head. Sections through the otic vesicle show that somitic muscle, located either side of the notochord in control embryos (c, arrowhead) is absent in β -APN treated embryos (j, arrowhead). A section through the abdomen of a β -APN treated embryo reveals that all the tissue types are present, although there is no evidence of gut coiling (compare d & k). Sections through the tail show that the notochord is malformed in β -APN treated embryos (compare e, f & g with l, m & n). The notochord appears larger and bulges outward from its normal location beneath the neural tube. ov = otic vesicle; hrt = heart; nt = neural tube; no = notochord; so = somite

Figure 3.6



filled odemas, there was disruption to the somites, which appeared loosely packed, and normal coiling of the gut had not taken place (compare Figure 3.5I & J). Sections from different areas along the length of β -APN treated embryos revealed defects throughout the A-P axis (Figure 3.6). Embryos cultured until stage 43/44 in the presence of 0.25mM β -APN displayed an abnormal shape to the head, a malformed gut, and bulges of the notochord due to its kinking along the full-length of the tail (Figure 3.6i-n). Additionally, there appeared to be an absence of muscle flanking the notochord at the level of the otic vesicles (see arrowhead in Figure 3.6c & j).

To further investigate the effect of β -APN treatment on both the somites and notochord, embryos were analysed with the monoclonal antibodies 12/101, to stain somites (Figure 3.7a-g), and MZ-15, to stain the notochord (Figure 3.7h-j). At stage 33/34 the muscle in β -APN treated embryos appears identical to the control (Figure 3.7a & b), but by stage 37/38 the muscle has shifted to a more posterior position relative to the control (Figure 3.7c & d, arrowheads). At stage 44/45 this posterior shift is clearly visible with the most anterior muscle blocks pushed well behind the otic vesicle (Figure 3.7e-f, arrowheads). MZ-15 staining indicates that the notochord is morphologically normal at stage 36 (Figure 3.7h), but at stage 45 a number of kinks have developed in the tail (Figure 3.7i & j). Since MZ-15 stains undifferentiated notochord only the posterior tip of the notochord is stained in Figure 3.7j, however kinks were visible along the length notochord. Staining of the muscle and notochord indicated that both structures form normally in the presence of β -APN and that the posterior shift of the somites and kinking of the notochord are late effects.

Figure 3.7 Immunolocalisation of muscle and notochord markers in embryos treated with β -APN

Control (a, c, e) and 0.25 mM β -APN treated (b, d, f) embryos were collected at stage 33/34 (a, b), stage 37/38 (c, d), and stage 44/45 (e, f). They were then probed with either muscle (12/101) or notochord (MZ15) specific antibodies, although the latter also detects the otic vesicle. (a-g) Up to, and including, stage 33/34 the somitic muscle of β -APN treated embryos appears to have formed normally (compare a & b), but by stage 37/38 it has moved posteriorly when compared to control embryos (c, d, arrowheads). This posterior shift of the somitic muscle is more evident at stage 44/45 (e, f, g). An enlarged view of the head and anterior abdomen of a stage 44/45 β -APN treated embryo shows the large gap separating the anterior most somites (as) and the otic vesicle (ov) (arrowheads). At stage 33/34, the notochord appears to be normal in β -APN treated embryos (h) were also collected and the notochord visualised using the MZ-15 antibody (h-j). At stage 33/34 the notochord appears to have formed normally (h), but at stage 44/45 the notochord is severely kinked (i), as made evident by an enlarged view of the tip of the tail. ov = otic vesicle; as = Anterior most somite.

Figure 3.7

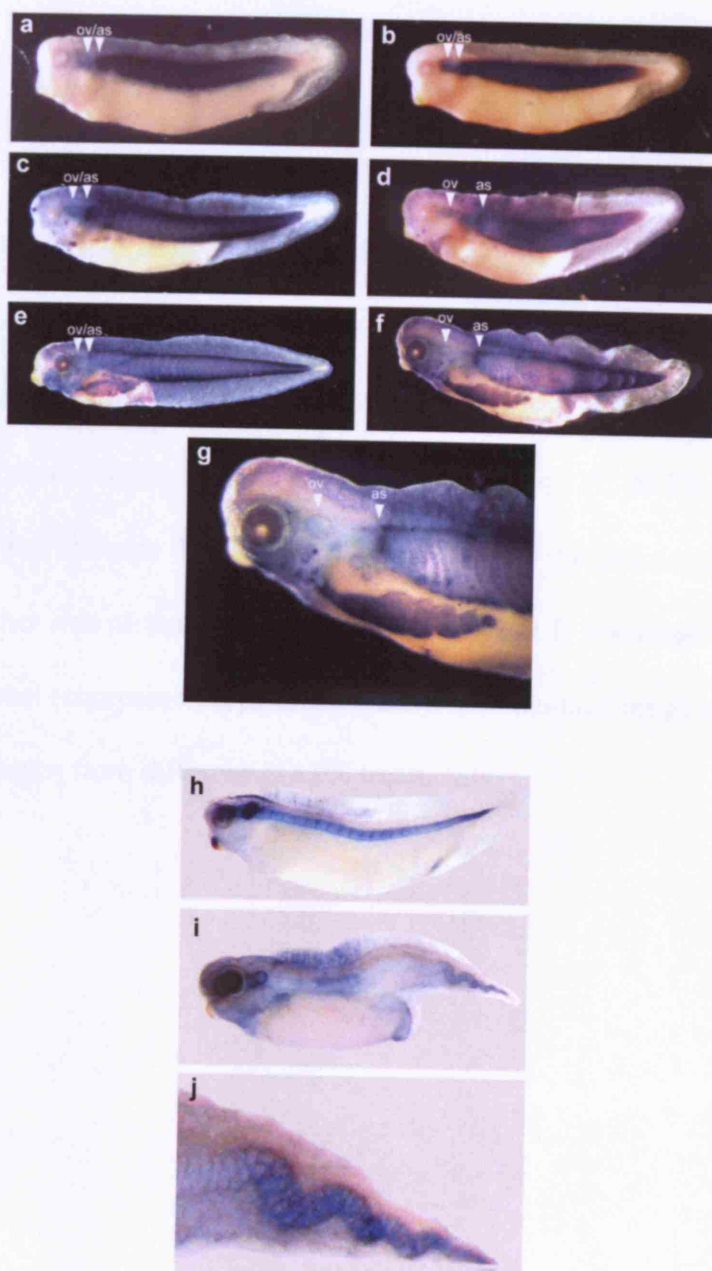
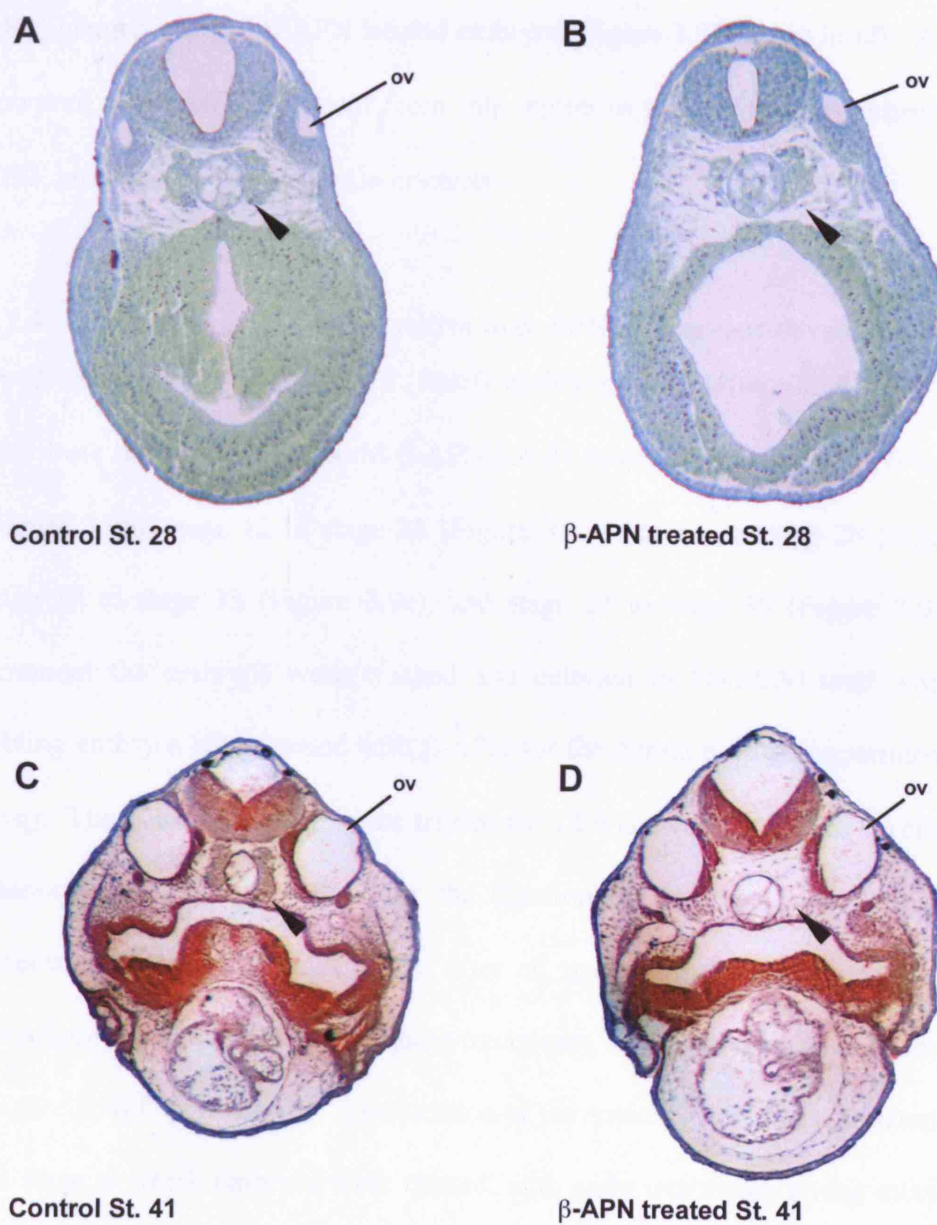


Figure 3.8 Muscle is present in anterior regions of β -APN treated embryos at stage 28 but absent at stage 41

Embryos cultured in either NAM/20 (A & C) or NAM/20 containing 0.25mM β -APN (B & D) were collected at stage 28 (A & B) or 41 (C & D), stained and sagittal sections taken through the embryo. At the level of the otic vesicle (ov) muscle was present either side of the notochord at stage 28 (A & B arrowheads) where as at stage 41 it is absent (compare C & D arrowheads). An identical result was obtained with at least 4 embryos from different β -APN treatments.

Figure 3.8



To confirm that muscle is present earlier in development, sagittal sections were taken through an anterior region of the embryo at the level of the otic vesicle. The otic vesicle was used as an easily identifiable structure both early and late in development. This revealed that at stage 28 muscle flanks the notochord in both control (Figure 3.8A, arrowhead) and β -APN treated embryos (Figure 3.8B, arrowhead). At stage 41 however, the muscle is absent from this region in β -APN treated embryos (Figure 3.8D, arrowhead), but present in controls.

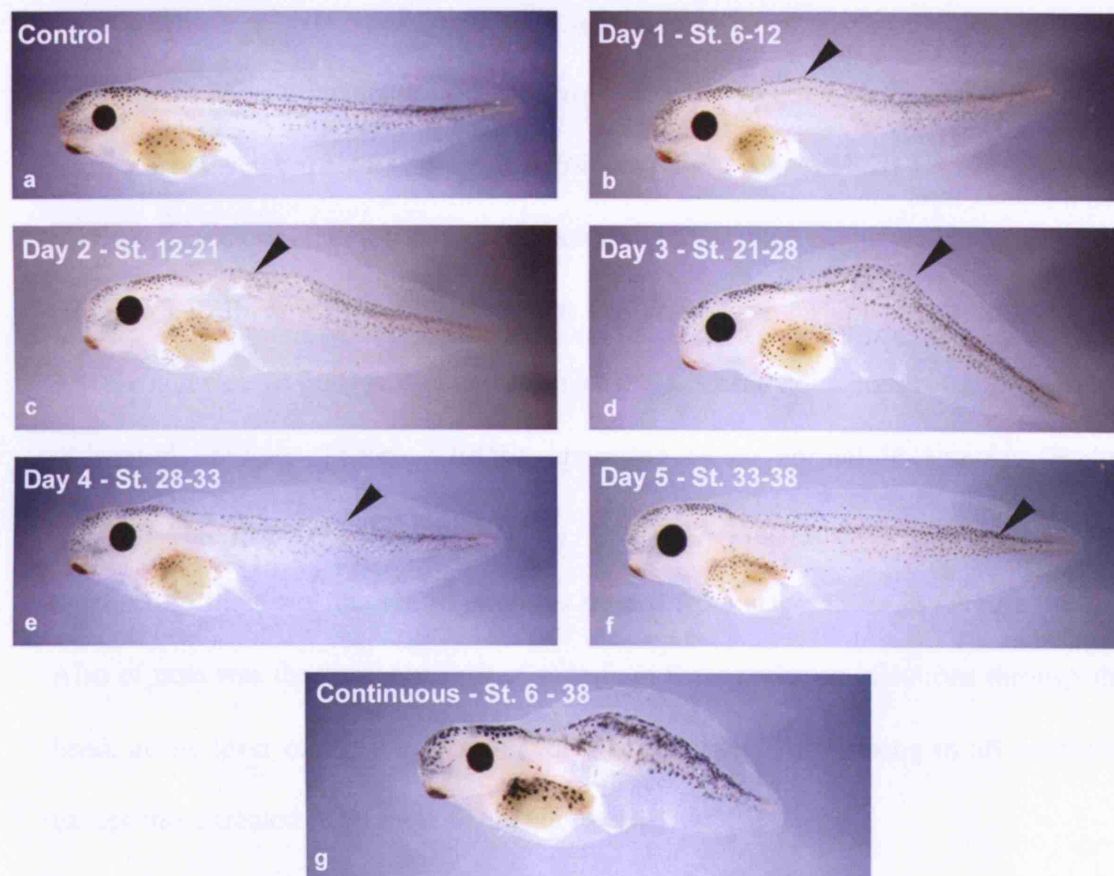
3.2.4 *X. laevis* embryos are sensitive to β -APN throughout development

To determine the stages when *X. laevis* embryos are sensitive to β -APN treatment they were incubated in 0.25mM β -APN for 24 hours only, from stage 6 to stage 12 (Figure 3.9b), stage 12 to stage 21 (Figure 3c), stage 21 to stage 28 (Figure 3.9d), stage 28 to stage 33 (Figure 3.9e), and stage 33 to stage 38 (Figure 3.9f). After treatment the embryos were washed and cultured in NAM/20 until stage 41/42. Sibling embryos were treated with β -APN for the duration of the experiment (Figure 3.9g). The phenotype of embryos treated for 24 hours was much less severe than the phenotype of embryos treated for the duration of the experiment, but notochord defects were nevertheless observed after all treatments. Although kinking of the notochord was apparent after all these treatments, it was only found in a small region (Figure 3.9b-f, arrowheads). The location of the notochord defects was determined by the stage at which embryos were treated, with early treatments giving more anterior defects than later treatments. Embryos treated continuously for the duration of the experiment exhibited defects along the whole length of the notochord (Figure 3.9g). The progressive movement of notochord defects from

Figure 3.9 β -APN treatment during different stages of development

Embryos were treated with 0.25mM β -APN between stage 6 and stage 12 (b), stage 12 and stage 21 (c), stage 21 and stage 28 (d), stage 28 and stage 33 (e), and stage 33 and stage 38 (f). In addition, some sibling embryos were incubated in β -APN for the duration of the experiment, between stage 6 and stage 38 (g). Sibling control embryos were never treated with β -APN (a). Kinks in the notochord, and surrounding tissue, appeared progressively more posterior the later in development β -APN was applied (arrowheads). This result was obtained in 100% of embryos treated and was reproducible in two independent experiments (n=>50).

Figure 3.9



anterior to posterior was confirmed by sectioning these embryos through the tail (Figure 3.10). In embryos treated from stage 6-12 the notochord is both enlarged (in diameter) and deformed in the most anterior section (Figure 3.10A), before returning to normal morphology posteriorly (Figure 3.10Aa-c). For sections through the tail of control embryos refer back to Figure 3.6e-g, which were obtained in the same experiment. A similar disruption is seen in anterior sections of the tail of embryos treated from stage 12-21 (Figure 3.10Ba) and from stage 21-28 (Figure 3.10Ca). However, in these embryos notochord defects extend into middle sections of the tail (Figure 3.10Bb & Cb), while the most posterior sections appear normal (Figure 3.10Bc and Cc). In embryos treated from stage 28-33 the notochord is only deformed in central sections (Figure 3.10Db), appearing to be normal in anterior (Figure 3.10Da) and posterior (Figure 3.10Dc) sections. Finally, notochord defects are localised to the tip of the tail in embryos treated from stage 33 to 38 (Figure 3.10E). Also of note was the posterior shift of muscle in these embryos. Sections through the head, at the level of the otic vesicle, show that muscle was missing in all embryos, except those treated from stage 6-12 (Figure 3.10Fa-f).

Figure 3.10 Notochord defects appear more posteriorly the later in development β -APN is applied

Embryos from the experiment presented in Figure 3.9 were sectioned and examined histologically. Representative sections through the tail (A-E) and head (F), for each time point, are presented. Embryos incubated in β -APN between stage 6 and 12 had both an enlarged and deformed notochord in the most anterior section of the tail presented (Aa), returning to a normal morphology in more posterior sections (Ab-c). Notochord defects are also found in anterior tail sections of embryos treated between stage 12-21 and stage 21-28, but these defects now extend into mid-tail sections (Ba & b) and (Ca & b). Mid-tail sections are the only ones with an enlarged and deformed notochord in embryos treated between stage 28-33 (Db). Finally, in embryos treated between stage 33 to 38 only the most posterior section in the notochord enlarged and deformed (Ec). These embryos should be compared with the control embryo in Figure 3.6, which came from the same fertilisation. Panel F shows sections through the head of these β -APN treated embryos and shows the absence of muscle flanking the notochord at the otic vesicle level occurs at most time points, the exception being embryos treated between stage 6-12 (Fb arrowhead). Control (Fa), stage 6-12 (Fb), stage 12-21 (Fc), stage 21-28 (Fd), stage 28-33 (Fe) and stage 33-38 (Ff).

Figure 3.10

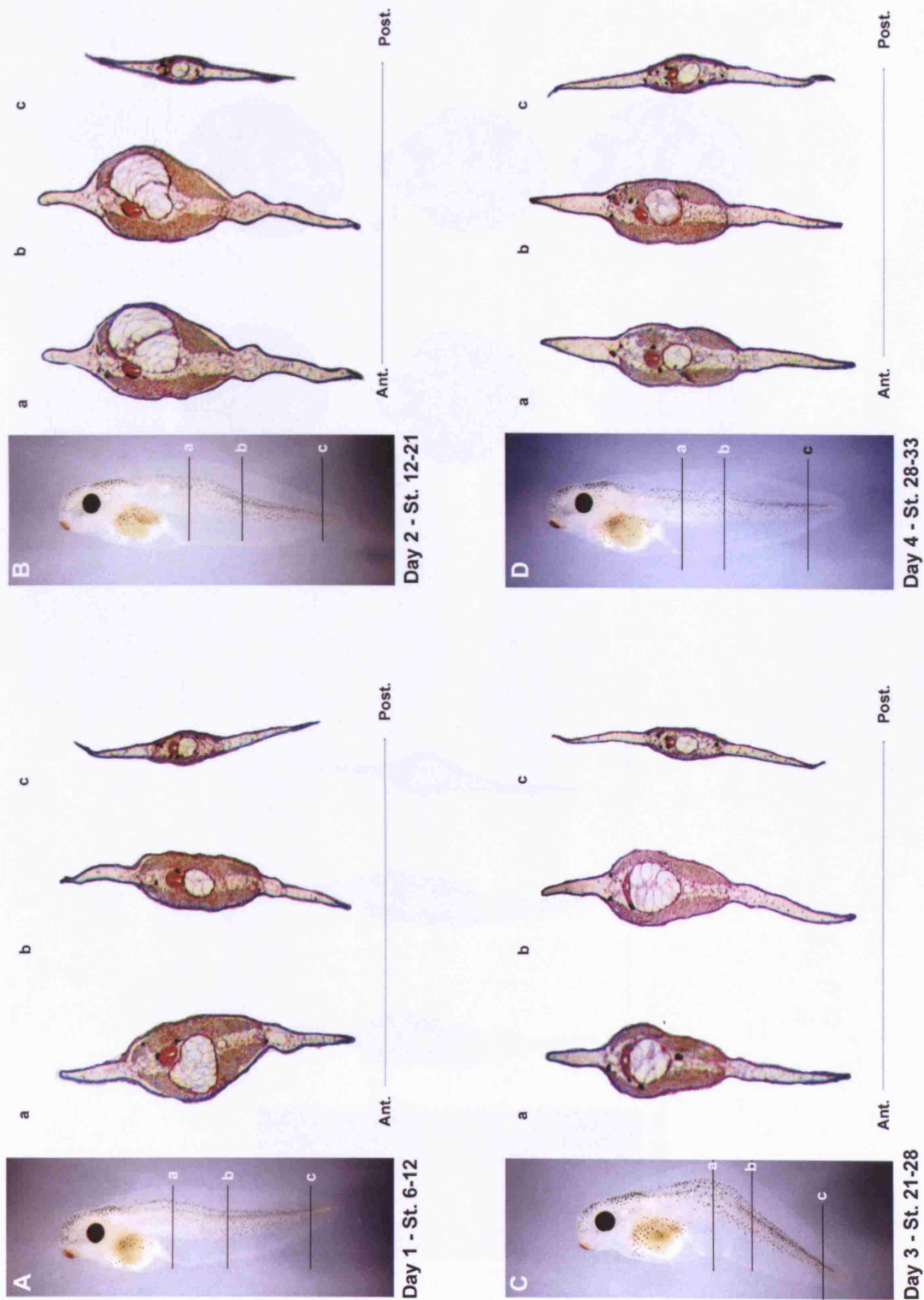
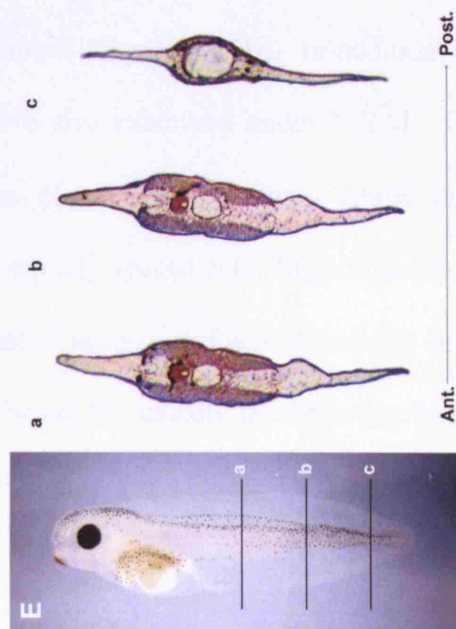
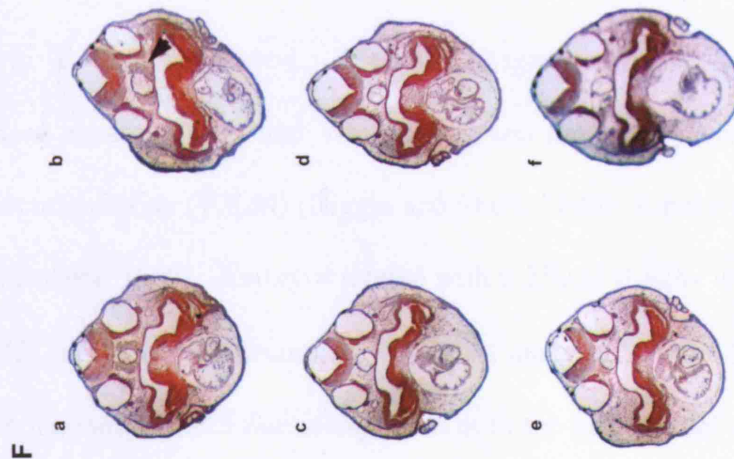


Figure 3.10 cont.



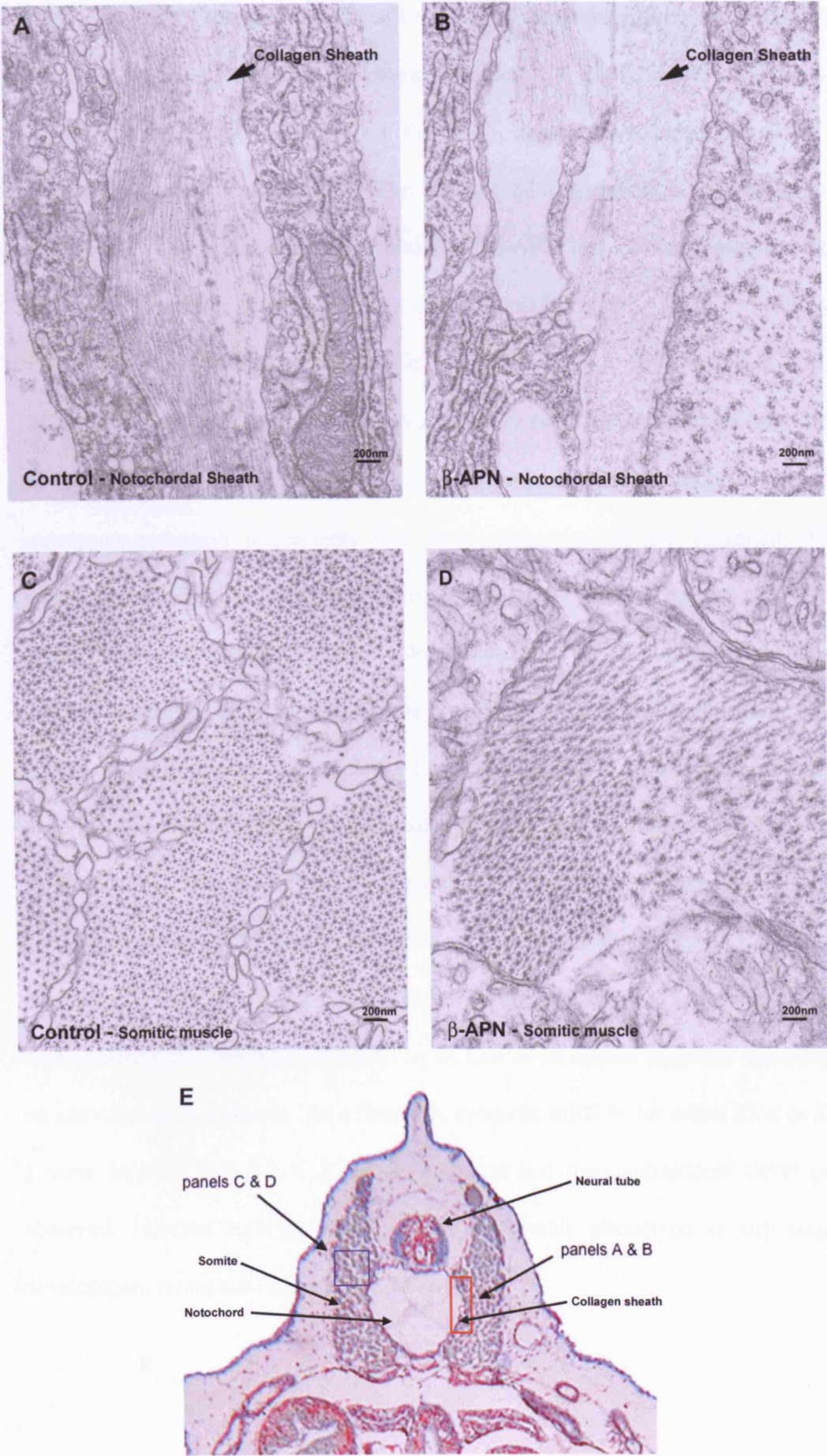
3.2.5 The notochordal sheath and somitic muscle is severely disrupted in β -APN treated embryos.

Since Lox is required for collagen fibre formation in mammals (Hornstra et al., 2003; Maki et al., 2002), we anticipated that β -APN would disrupt the formation of these fibres in *X. laevis* embryos. Type II collagen fibres are very abundant in the notochord sheath (Smith and Watt, 1985) and are clearly visible by transmission electron microscopy (T.E.M) (Riggin and Shultz, 1986; Schultz et al., 1985; Snawder and Chambers, 1990). Embryos treated with 0.25mM β -APN were therefore fixed at stage 42, processed and sectioned for T.E.M analysis. Figure 3.11A and B show a 20000x magnification of the collagen fibres in the notochordal sheath. In transverse sections through control embryos, collagen fibres run longitudinally around the notochord in long uninterrupted fibres that form the sheath (Figure 3.11A). Collagen fibres are also visible in the notochord sheath of β -APN treated embryos, but there are very few long fibres and little cross-linking between individual collagens when compared to the control (Figure 3.11B). In addition, transverse sections through the somitic muscle were also examined under T.E.M. Control muscle contains highly regular formations of myosin and actin fibres (a single central myosin fibre surrounded by six equally spaced actin fibres), in cross-section (Figure 3.11C). This regular arrangement of myosin and actin fibres can be seen in some sections through β -APN treated embryos, but usually the fibres are less distinct and are frequently cut in an almost longitudinal plane. (Figure 3.11D).

Figure 3.11 T.E.M analysis of collagen fibres in the notochordal sheath and somitic muscle in β -APN treated embryos

Embryos treated with 0.25mM β -APN were processed for T.E.M analysis and observed at 20000x magnification. In untreated control embryos (A), collagen runs in long relatively uninterrupted fibres longitudinally around the notochord. In embryos treated with β -APN (B) the fibres are considerably disorganised and have not formed the long fibres found in untreated controls. The notochord lies to the left of the sheath but is not shown in each photograph. The somitic muscle of control embryos (C) is highly ordered with single myosin fibres surrounded by six actin fibres running anterior to posterior through the muscle. A similar organisation can be seen in some sections of β -APN treated embryos, but the fibres frequently appear to move out of plane within the section and do not form the regular structures found in untreated embryos (D). The approximate positions of each section are shown in the diagram (E). I gratefully acknowledge Mark Turmaine (UCL) for embedding and sectioning these embryos for T.E.M analysis.

Figure 3.11



3.2.6 Injecting *Xlox* and *Xloxl-1* mRNA has no effect on embryonic patterning

Since *Xloxl-1* is expressed in the dorsal mesoderm of gastrulae we considered the possibility that it was necessary for the cell rearrangements characteristic of this important stage of development. The absence of a gastrulation phenotype in the whole embryo embryos could be explained if β -APN had difficulty penetrating the embryo at this stage. To overcome this potential problem, DMZ explants were isolated from early gastrulae (stage 10) and incubated in 0.25 mM β -APN (Figure 3.12). It was anticipated that β -APN would easily penetrate these fragments through the open wound edge. When cultured in isolation, DMZ fragments undergo convergent-extension movements that drive elongation of the fragment (Figure 3.12A) and differentiate notochord, muscle, and neural tissue (Figure 3.12C). DMZ fragments incubated in β -APN also undergo elongation that is indistinguishable from controls (Figure 3.12B) and differentiate notochord, muscle, and neural tissue (Figure 3.12D). As a further test for a role for Lox-like enzymes in gastrulation, animal caps were isolated from blastulae and incubated in the mesoderm inducing factor activin. Animal caps treated in this way undergo convergent-extension movements that drive a pronounced elongation of the cap during gastrula stages (Figure 3.13C). Elongation was not inhibited by adding β -APN to the incubation media (Figure 3.13D). This data confirms that inhibiting the activity of Lox or its related enzymes has no effect on gastrulation in *X. laevis*. As a final test, synthetic mRNAs for either *Xlox* or *Xloxl-1* were injected into 2-cell *X. laevis* embryos and their subsequent development observed. Injected embryos displayed no discernible phenotype at any stage of development (compare Figure 3.14B&C with A).

Figure 3.12 β -APN does not prevent the elongation of isolated DMZ's, nor the formation of muscle and notochord in these explants

DMZ's were isolated from stage 10/10.5 embryos and cultured until sibling control embryos had reached stage 25, in the presence (B) and absence (A) of 0.25mM β -APN. Elongation of the DMZ's occurred in both sets of explants and histological analysis revealed that both notochord (no) and muscle (mu) were present (C & D arrowheads)

Figure 3.13 β -APN does not inhibit elongation of activin treated animal caps

Animal caps were removed from stage 8/9 blastulae and incubated until sibling control embryos had reached stage 25. (A) Untreated controls. (B) 0.25mM β -APN. (C) Activin protein. (D) Activin protein plus 0.25mM β -APN (D). Activin caused the elongation of treated caps (C), which was not inhibited by the presence of β -APN (D).

Figure 3.12

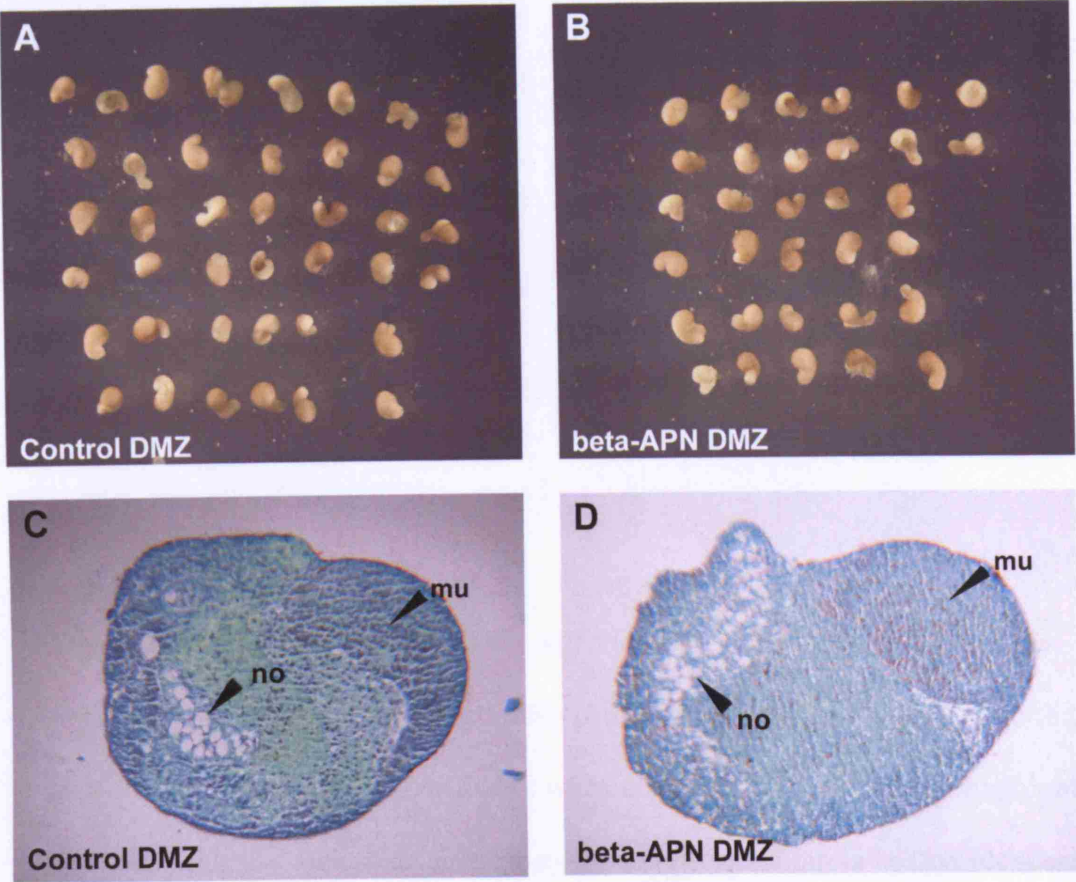


Figure 3.13

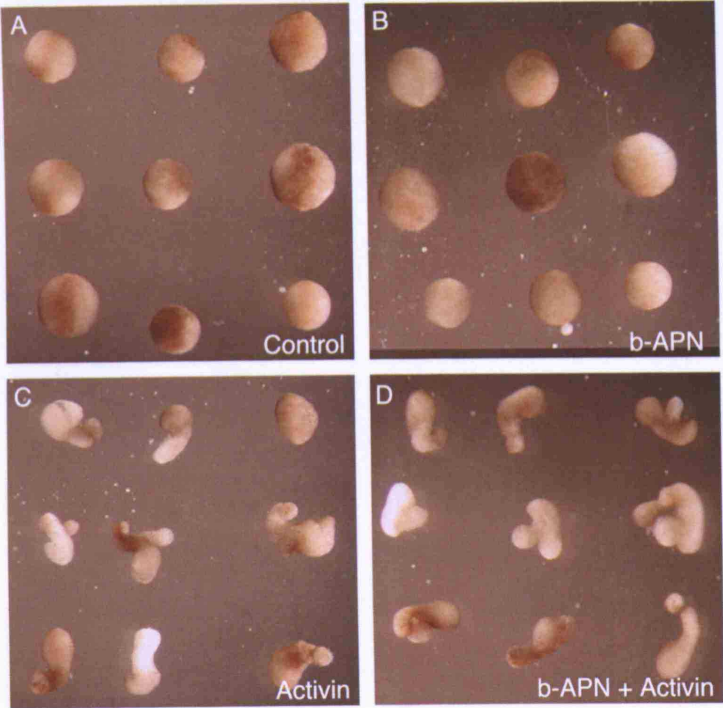
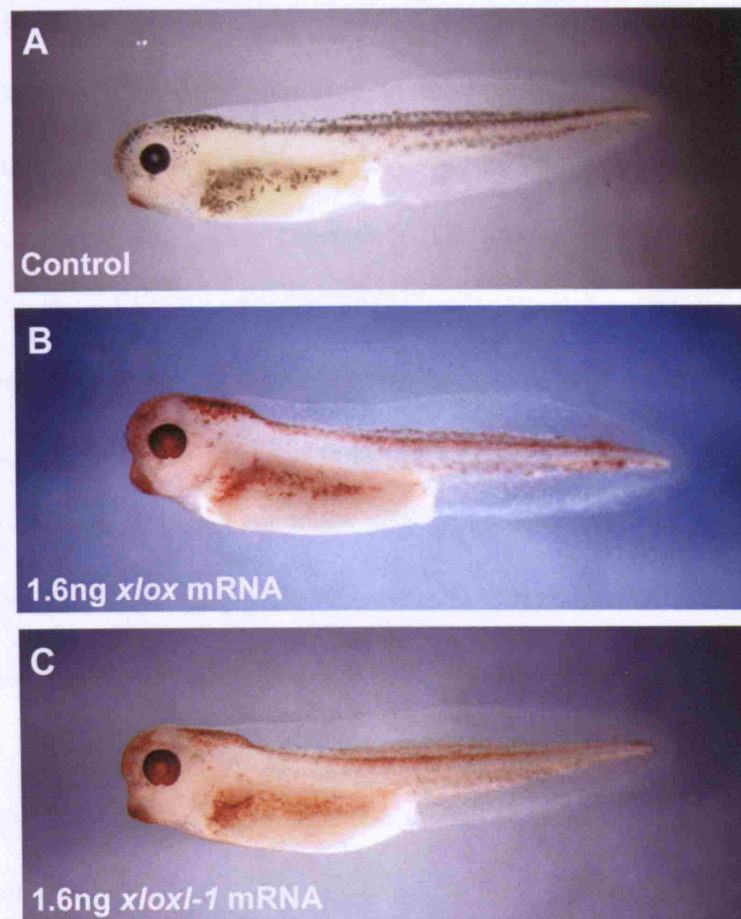


Figure 3.14 Over-expression of *Xlox* and *Xloxl-1*

Injection of 1.6ng synthetic capped mRNA for either *Xlox* (B) or *Xloxl-1* (C) did not generate any visible phenotypic defect when compared to uninjected sibling controls (A). All embryos appeared identical to the uninjected controls in two independent sets of injections.

Figure 3.14



3.3 Discussion

3.3.1 *X. laevis* lox genes

Despite previous use of *X. laevis* embryos to study the toxicology of specific Lox inhibitors (Dawson, 1993; Dawson et al., 2000; Dawson et al., 2002; Dawson et al., 1990; Snawder and Chambers, 1993), neither Lox nor any other members of the LOX family have previously been identified in this organism. In this chapter I have identified three members of the Lox family that are expressed during early *X. laevis* embryonic development. *Xlox* encodes a 429 amino acid protein with a highly conserved C terminal LOX catalytic domain, while *Xloxl-1* encodes a 528 amino acid protein with a conserved LOX domain and additional proline rich domain between amino acids 193 and 294. This proline rich domain appears to account for the 99 amino acid difference between *Xlox* and *Xloxl-1*. Also described is a partial sequence for *Xloxl-3* encoding the C-terminal 305 amino acids of the full-length protein. Mammalian *Loxl-3* contains four C-terminal SRCR domains thought to be involved in the localisation of the protein to cell membranes (Sarrias et al., 2004). Only the last 71 amino acids of the fourth SRCR domain is included in *Xloxl-3* described here.

Lox is secreted as an inactive proenzyme that is activated by Procollagen-C-proteinase (PCP) (Panchenko et al., 1996), an enzyme that removes the C-terminal pro-peptide of fibrillar collagens and was shown to be the same enzyme as BMP-1 (Kessler et al., 1996; Li et al., 1996). The cleavage site for BMP-1/PCP has been conserved in the *X. laevis* sequence presented here indicating that XBMP-1 is likely to activate XLox. In contrast, BMP1/PCP cleavage sites identified in mammalian *Loxl-1* have not been conserved in XLoxl-1, despite *in vitro* studies on the mammalian

protein suggesting it is activated by BMP-1 (Borel et al., 2001). Activation of mammalian Loxl-1 by BMP-1/PCP is either an artefact of the *in vitro* studies, *in vivo* cleavage sites required for activity have yet to be described, or Xloxl-1 is regulated differently from its mammalian orthologues. Clarification of this point will require further studies. XLoxl-1 does include a conserved tribasic (RRR), furin-like, cleavage site used to produce an inactive proLoxl-1 in mammals, indicating some conservation of the mechanisms responsible for regulating the activity of this enzyme.

3.3.2 Function of *lox* genes in early *X. laevis* development

Xloxl-1 is maternally expressed, while *Xlox* and *Xloxl-3* are first detected at stage 13/14 by RT-PCR. Using whole mount *in situ* hybridisation I have shown that transcripts for all three genes are localised to the notochord. Interestingly, *Xloxl-1* transcripts are found in the dorsal mesoderm of early gastrulae, in the presumptive notochord. This precedes the expression of the fibrillar collagens, which first appear during early neurulation with the expression of *proCol2 α 1* (see Figure 3.3) (Su et al., 1991), and indicates that Xloxl-1 might act on other, as yet unidentified, substrates. XLoxl-1 expressing cells reside in a region undergoing convergent extension movements (Keller et al., 2003), suggesting that XLoxl-1, or its substrates, might play a role in this process. In support of this suggestion, sea urchin embryos incubated in β -APN, a specific inhibitor of Lox-related enzymes, fail to gastrulate (Butler et al., 1987; Wessel and McClay, 1987), while *Lox* is upregulated in invasive metastatic breast cancers (Kirschmann et al., 2002). However, *X. laevis* embryos gastrulate normally and both DMZ and activin induced animal cap explants still undergo elongation as a result of convergent extension movements in the presence of β -APN.

This indicates that *lox* genes are not required for the convergent extension movements of gastrulation in *X. laevis*.

3.3.3 *X. laevis* Lox function in notochord and somite development

All three *lox* genes are predominantly localised to the notochord, a stiff rod-like structure running the length of the embryo from the mid-brain to the tip of the tail, that is surrounded by a sheath rich in collagen (Stemple, 2005). The notochord provides tensile strength to the A-P axis prior to the development of skeletal structures, as well as acting in the capacity of a signalling centre conveying positional information to the endoderm, neural tube, and somites. Fibrillar collagens and elastin are the best characterised substrates of Lox. In *X. laevis* *proCol2a1* transcripts are localised to the notochord, somites and head (Bieker and Yazdani-Buicky, 1992; Seufert et al., 1994; Su et al., 1991), as are *proCol5a1*, *proCol5a2*, and *proColl1a1* (see Chapter 6). The distribution of these collagens is therefore consistent with the position of *Xlox* expression, and its potential function for collagen fibre formation in the notochord. Substrates for Loxl-1 have yet to be identified, however, murine Loxl-1 does interact with fibulin-5, an elastin associated protein, and localises to sites of elastin expression (Liu et al., 2004). Unfortunately *X. laevis* homologues for these two genes have yet to be described and we cannot compare the distribution of their transcripts to those for *Xloxl-1*. However, results from studies in chick embryos find *elastin* transcripts are localised to the notochord and somites from 23hrs of development, which is roughly equivalent to stage 16/17 in *Xenopus* (Visconti et al., 2003). This is at the same time, and in the same structures that *Xlox* and *Xloxl-1* are present in early *Xenopus* embryos suggesting that one or both might interact with a homologue of elastin.

3.3.4 Inhibition of LOX function using β -APN

Addition of β -APN, a specific nitrile inhibitor of the Lox catalytic domain (Tang et al., 1983), to *X. laevis* embryos resulted in numerous defects, most notably in the notochord and somites but also poor gut differentiation and development of blisters in the epidermis. Although we have not shown that β -APN inhibits LOX activity in *X. laevis* embryos we believe that this is the case. β -APN has been shown to inhibit several members of the LOX family in mammals, including Lox (Tang et al., 1983) and Loxl-1 (Borel et al., 2001). We have shown that both enzymes are highly conserved between mammals and *X. laevis*, suggesting that the *Xenopus* enzymes will also be inhibited by β -APN. In addition, β -APN has been shown to inhibit Lox related enzymes in both *Drosophila melanogaster* (Molnar et al., 2005) and sea urchins (Butler et al., 1987). This demonstrates that β -APN is an inhibitor of LOX activity in both vertebrates and invertebrates, adding further weight to our belief that it inhibits *Xenopus* LOX activity.

From the known substrates of Lox we considered the possibility that the observed defects were due to disruption of collagen and elastin fibre formation. Transcripts for many fibrillar collagens are highly enriched in the notochord of *Xenopus* embryos and to a lesser extent in the adjacent somites (Su et al., 1991) (Chapter 6, this thesis). Type II collagen protein is also abundant in the notochord sheath (Smith and Watt, 1985). Although *X. laevis* elastin has yet to be described, it has been shown to be abundant in the notochord sheath of the chick embryo (Visconti et al., 2003). We therefore used T.E.M sections to analyse the notochord sheath and showed that the collagen fibres were indeed disrupted in β -APN treated embryos. In control embryos,

collagen runs around the notochord in long uninterrupted fibres, while in β -APN treated embryos there are very few long fibres and little cross-linking between individual collagen molecules.

The notochord sheath is believed to have at least two functions that could provide an explanation for the abnormalities observed in β -APN treated embryos (Stemple, 2005). First, the sheath resists the tendency of notochord cells to osmotically swell, causing the internal pressure of these cells rise. Therefore the notochord becomes a stiff rod. If the sheath is enzymatically removed then the notochord lengthens more rapidly than adjacent tissues, such as the neural tube and paraxial mesoderm, and buckles and folds (Adams et al., 1990). The disrupted sheath of β -APN treated embryos could also allow the notochord to lengthen rapidly, providing an explanation for the kinked notochord. Second, the notochord acts as an axial skeleton to which the skeletal muscles of the somites are attached via the sheath. The disrupted sheath of β -APN treated embryos may not be sufficiently strong to withstand the strains of swimming, giving way as the tailbud stage embryo becomes increasingly active. Therefore the somites would slip towards the tail disrupting the morphology of the notochord in posterior regions. This could explain why the notochord defects are not observed until about stage 40, 2-3 days after the notochord begins to elongate in late neurula stages.

Disrupting collagen and elastin cross-linking in the ECM may also explain the other phenotypes observed in β -APN treated embryos, including blistering of the skin and

the formation of oedemas. Both could be accounted for by disrupting the connective tissue of the skin and the spaces between tissues. It may also explain the poor differentiation of the gut since interactions between the endodermal layer of the gut and its surrounding mesoderm, such as connective tissue and smooth muscle, are known to be required for gut differentiation (Horb and Slack, 2001; Okada, 1957; Okada, 1960).

3.3.5 Summary

In this chapter I have identified three members of the LOX family of enzymes, all of which are expressed in the notochord of the developing embryo. Disruption to their function by specific inhibitors of the catalytic domain results in defects in collagen formation that consequently disrupts the notochord and somitic muscle. Despite indications that XLoxl-1 might play a role in dorsal-ventral patterning, this was found not to be the case. This work has been accepted for publication in *Differentiation* (Geach and Dale, 2005).

Chapter 4

The first and second CUB domains are required for Xld to process Chd

4.1 Introduction

4.1.1 Tld metalloprotease activity

The Tlds consist of an N-terminal astacin-like protease domain linked to a variable number of C-terminal CUB and EGF-like domains. For example, BMP-1, which arises due to alternative splicing of the mammalian *tld* (*mtld*) gene, contains three CUB domains and a single EGF domain, with the arrangement C-C-E-C (Wozney et al., 1988). In contrast, mTld contains five CUB domains and two EGF domains, with the arrangement C-C-E-C-E-C-C (Takahara et al., 1994b). This latter arrangement is shared by most other members of the Tld family, including mammalian Tll-1 and Tll-2 (Scott et al., 1999), *Xenopus* Xld and Xlr (Dale et al., 2002; Goodman et al., 1998), *Zebrafish* Tld (Blader et al., 1997), and *Drosophila* Tld (Shimell et al., 1991). Intriguingly, whereas BMP-1 cleaves the dorsalising protein Chd (Scott et al., 1999; Wardle et al., 1999) mTld does not (Scott et al., 1999), despite arising from the same gene. BMP-1 and mTld have an identical protease domain, which indicates that the arrangement of C-terminal CUB and EGF domains is responsible for the difference in this activity. However, Tll-2, Xld, Xlr, and zTld all have an identical domain structure to that of Tld (Figure 4.1), yet all four proteases cleave Chd (Blader et al., 1997; Dale et al., 2002; Piccolo et al., 1997; Scott et al., 1999). This indicates the difference between activities of BMP-1 and mTld is not simply the addition of extra CUB and EGF domains to the latter.

4.1.2 CUB domains

CUB domains are approximately 110 amino acids long and usually contain four cysteine residues that form two disulphide bridges (Bork and Beckmann, 1993). Because this module was originally found in the mammalian complement proteases C1s and C1r (C), the sea urchin ECM protein Uegf (U), and the mammalian metalloprotease BMP-1 (B), the term CUB was adopted (Bork and Beckmann, 1993). The module was subsequently found in many other extracellular proteins, including the proteases MASP2 and RaRF (Takayama et al., 1999; Thiel et al., 1997), sea urchin SpAN/BP10 (Lepage et al., 1992; Reynolds et al., 1992), some members of the ADAMTS family (Zheng et al., 2001), and various vertebrate hatching enzymes (Fan and Katagiri, 2001; Geier and Zwilling, 1998; Yasumasu et al., 2005). CUB domains have also been found in procollagen C-proteinase enhancer (PCPE) (Takahara et al., 1994a), the mammalian cell adhesion molecule neuropilin (Giger et al., 1998), the mammalian hyaluronate-binding protein TSG-6 (Wisniewski et al., 1994), mammalian spermadhesins (Menendez et al., 1995), and the sea urchin ECM proteins fibropellin I and III (Delgadillo-Reynoso et al., 1989). Most of the available evidence suggests that CUB domains are involved in protein-protein interactions. For example, CUB domains mediate dimer formation for both C1r and C1s and their interaction with C1q, generating the first component (C1) of the classical pathway of the mammalian complement system (Busby and Ingham, 1988). In the lectin complement pathway, MASP-2 binds the mannose binding protein (MBP) via two CUB domains and a single EGF-like domain in a Ca^{2+} dependent manner (Feinberg et al., 2003; Wallis and Dodd, 2000). Additionally, the PCPE protein, consisting of two CUB domains and a non-homologous domain, binds to both BMP-1 and procollagen type I (proCol1I) in turn enhancing the processing of proCol1 α 1 into a mature

collagen fibril (Adar et al., 1986; Takahara et al., 1994a). Two naturally occurring isoforms of PCPE, which consist of only the two CUB domains, retain the enhancer activity (Kessler and Adar, 1989), indicating that the binding capacity of PCPE resides in these modules.

In *Xenopus*, complete removal of the protease domain of BMP-1 and Xld, leaving just the signal sequence and CUB domains, results in a dominant negative version (dnBMP-1 and dnXld) of these proteins (Wardle et al., 1999). These dorsalise ventral mesoderm in VMZ assays, an effect most likely due to their binding Chd and protecting it from cleavage by the endogenous protease. DnBMP-1 and dnXld block cleavage of Chd by both BMP-1 and Xld, suggesting some functional redundancy of the CUB domains between these two proteases. Additionally, the introduction of a point mutation within the active site of Xld (Tyr²⁷²-Asn) both blocks the proteolytic activity of Xld and generates a dominant-negative protein, indicating a role for the CUB domains in interactions with Chd (Piccolo et al., 1997). Despite indications that the CUB domains of the Tlds are involved in substrate binding, little is understood as to which domains are involved in this process. Hartigan et al., (2003) have recently shown that the 1st CUB domain of BMP-1 is required for secretion of the protease and the 2nd CUB domain for processing of procollagen type I (proCol1 α 1). Mammalian Tld also cleaves proCol1 α 1 but only the 1st, 2nd and 3rd CUB domains and 1st EGF-like domain are required to bind the substrate (Sieron et al., 2000), further indicating that the 4th and 5th CUB domains of mTld are functionally redundant.

4.1.3 EGF-like domains

Epidermal growth factor (EGF) is a signalling molecule derived from a precursor consisting of 8 EGF-like domains and the main EGF molecule itself. EGF-like domains can be separated into two subgroups; calcium-binding (cbEGF) and laminin-like EGF domains (Stenflo et al., 2000). All cbEGF domains contain 6 conserved cysteine residues, forming 3 disulphide bonds, and can be characterised by the consensus sequence CxCxxGFxxxxxxxxxC, along with an Asp/Asn hydroxylation site and a calcium binding sequence at their N-terminus (Prosite). Laminin-like EGF domains contain 8 cysteines forming 4 disulphide bonds. The Tld EGF-like domains are of the calcium binding flavour, although whether their capacity to bind Ca^{2+} is required for enzymatic activity remains to be seen. Three recent studies have shown that the EGF-like domains of BMP-1, mTld and mTll-2 are dispensable for cleavage of proColl α 1 and Chd (Garrigue-Antar et al., 2004; Hartigan et al., 2003; Petropoulou et al., 2005). In addition, deletion of the EGF-like domains of mTld actually allow Chd cleavage *in vitro*. Although not required for activity of the Tlds against any of the substrates so far investigated, they have so many substrates it is possible the EGF-like domains are required for some activities of the enzyme.

4.1.3 Summary

To examine the role of CUB and EGF domains in Xld, I have used PCR to generate constructs lacking a number of these domains. Removing successive CUB domains from the C-terminus I show that the 1st and 2nd domains are required for the cleavage of Chd and the second CUB domain is required for efficient secretion of the protease. However, the 2nd CUB domain alone is not sufficient for the protease to counteract the dorsalising activity of Chd. These results indicate that the 3rd, 4th and 5th CUB domains, as well as the two EGF-like domains, are functionally redundant in the protease for the purpose of Chd cleavage.

4.2 Results

4.2.1 Both CUB1 and CUB2 are required for the ventralising activity of Xld.

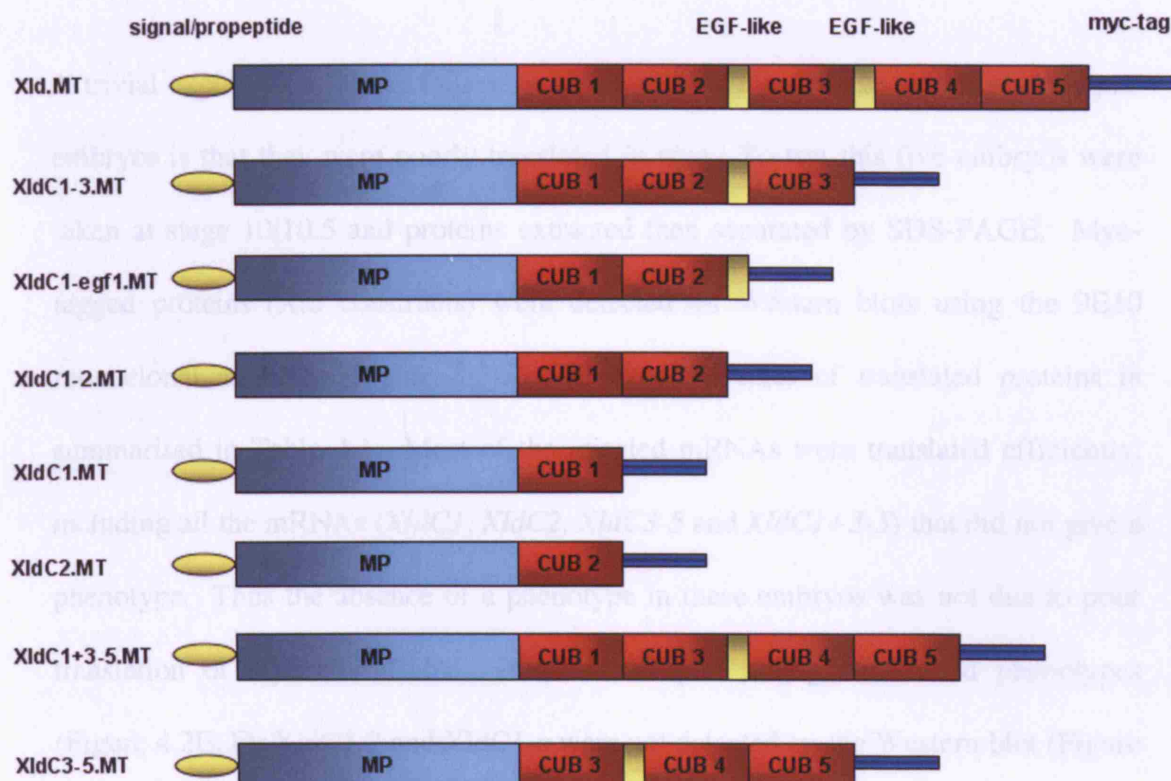
To investigate the role played by the C-terminal CUB and EGF-like domains in regulating the activity of Xld, PCR was used to remove specific domains from Xld containing 6-Myc tags at the C-terminus (Xld.MT). These tags allow us to detect the protease by Western blot analysis (Wardle et al., 1999) and do not adversely affect the proteolytic activity of Xld. Figure 4.1 illustrates the different deletion constructs produced and used in this study.

Injection of full-length *Xld* mRNA, such that the protease is expressed at increased levels throughout the developing *Xenopus* embryo, results in a reduction of dorsal-anterior structures, such as the head and notochord, and enlargement of posterior-ventral structures, such as the tail and ventral blood island (Goodman et al., 1998; Piccolo et al., 1997; Wardle et al., 1999). This is a result of mild mesodermal ventralisation and embryos displaying this phenotype will be described as ventralised in this study. A similar phenotype is observed when either *Xenopus* or zebrafish embryos are depleted of Chd (Oelgeschlager et al., 2003; Schulte-Merker et al., 1997), indicating that Xld is increasing cleavage of Chd. To determine the minimal domain structure of Xld required to ventralise *Xenopus* embryos, 800 pg of synthetic capped mRNA for each construct was injected into each blastomere at the two-cell stage and embryos cultured until uninjected siblings had reached stage 26/27 (Figure 4.2A). Injection either of *Xld* (Figure 4.2B), *XldC1-2* (Figure 4.2E), *XldC1-e* (Figure 4.2F), or *XldC1-3* (Figure 4.2G) mRNA resulted in a ventralised phenotype with an average DAI of 4, 3.5, 3.67 and 3.63 respectively

Figure 4.1 Schematic representation of CUB deletion constructs

PCR was used to introduce a number of domain deletions in the C-terminal region of a Myc tagged Xld construct (Xld.MT). Each construct included the signal/propeptide (yellow oval), the metalloprotease domain (MP), at least one CUB domain, and 6 Myc (9E10) epitopes at the C-terminus (small blue box) for detection in Western blot analysis.

Figure 4.1



(compare with Figure 4.2A). In contrast, injection either of *XldC1* (Figure 4.2C), *XldC2* (Figure 4.2D), *XldC3-5* (Figure 4.2H), or *XldC1+3-5* (Figure 4.2I) mRNA did not produce any phenotypic defects (DAI in all cases was 5). These results show that both the 1st and 2nd CUB domains are necessary for Xld to ventralise *Xenopus* embryos and that alone neither is sufficient.

A trivial explanation for the failure of some mRNA injections to ventralise *Xenopus* embryos is that they were poorly translated *in vivo*. To test this five embryos were taken at stage 10/10.5 and proteins extracted then separated by SDS-PAGE. Myc-tagged proteins (Xld constructs) were detected on Western blots using the 9E10 monoclonal antibody (Figure 4.3). The expected sizes of translated proteins is summarised in Table 4.1. Most of the injected mRNAs were translated efficiently, including all the mRNAs (*XldC1*, *XldC2*, *XldC3-5* and *XldC1+3-5*) that did not give a phenotype. Thus the absence of a phenotype in these embryos was not due to poor translation of injected mRNA. However, despite giving ventralised phenotypes (Figure 4.2E, F), *XldC1-2* and *XldC1-e* were not detected on the Western blot (Figure 4.3, lanes E and F). Subsequent sequencing of the plasmid revealed a single nucleotide frame shift between the last domain of the construct and the Myc epitope sequence, consequently preventing the Myc tag from being translated. *XldC1-2* and *XldC1-e* were therefore analysed in a rabbit reticulocyte lysate *in vitro* translation system to check that each mRNA was translated and produced a protein of the predicted size. Both *XldC1-2* and *XldC1-e* were translated efficiently *in vitro* (Figure 4.4A, B), producing proteins of 69 and 75 kDa respectively.

Figure 4.2 Ectopic expression of each deletion construct in whole embryos

Synthetic, capped, mRNA was made for *Xld* and each of the deletion constructs presented in Figure 4.1. 0.8ng of mRNA was injected into each blastomere at the two-cell stage. Embryos were left to develop until uninjected siblings reached stage 26/27 (A). Injection of *Xld* (B) resulted in mild ventralisation of the embryo (19/21 cases), characterised by a reduced head and enlarged posterior regions. This phenotype was reproduced by *XldC1-2* (23/23 cases) (E), *XldC1-e* (20/20 cases) (F), and *XldC1-3* (35/35 cases) (G). Injection of *XldC1* (C), *XldC2* (D), *XldC3-5* (H), and *XldC1+3-5* (I) did not produce any noticeable phenotype (0/25, 0/37, 0/32 and 0/31 ventralised cases respectively). The embryos shown are representative of the phenotypes observed.

Figure 4.3 *In vivo* translation of Xld deletion constructs

Five embryos injected with each Xld deletion, as described in Figure 4.2, were collected at stage 10/10.5, homogenised and subjected to western blot analysis with anti-Myc (9E10) antibody. With the exception of *XldC1-2* and *XldC1-e*, protein was detected for each construct and had the expected size. Extracts from *XldC2* injected embryos generally had four distinct bands, two of which migrated more slowly than expected. We have no explanation for this. One band migrated faster than expected and is probably the result of degradation. Sequencing of *XldC1-2* and *XldC1-e* revealed a frame shift at the junction with the Myc tag sequence, such that the resulting proteases would not be tagged.

Figure 4.4 *In vitro* translation of *XldC1-2* and *XldC1-e*

To confirm that mRNA for *XldC1-2* and *XldC1-e* was translated, and produced a protein of the predicted size, 0.5µg of mRNA for each construct was translated in an *in vitro* translation system incorporating S³⁵ labelled methionine to detect the product produced. Both mRNAs translated well in this reaction, producing proteins of 69 and 75 kDa respectively. This is the predicted size for these proteases lacking the C-terminal Myc tags.

Figure 4.2

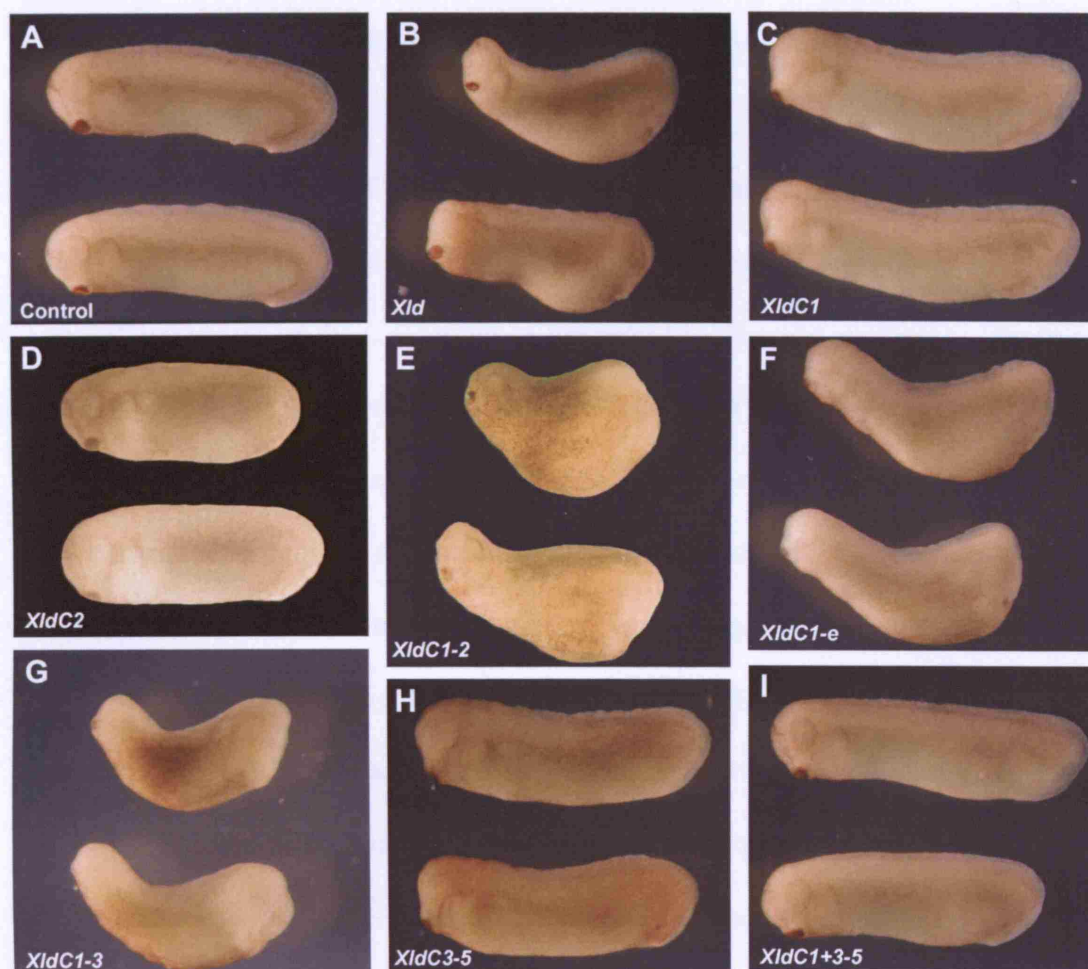


Figure 4.3

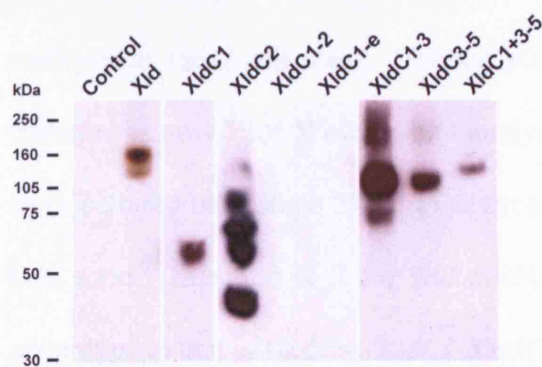
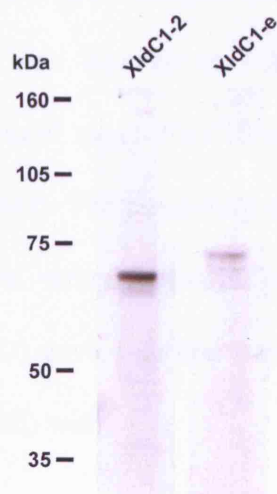


Figure 4.4



These were slightly smaller than the predicted sizes of both proteins as they lacked the 13 kDa Myc tag sequence. This suggests that the phenotypes generated by *XldC1-2* and *XldC1-e* (Figure 4.2D, E) were a result of the translation of these mRNAs.

4.2.2 Removal of both CUB4 and CUB5 does not enhance the activity of Xld

In vertebrates, BMP-1 and Tld are distinguished by the presence of additional CUB and EGF domains on the C-terminus of the latter (Takahara et al., 1994b). Since only BMP-1, and not mTld, has chordinase activity (Scott et al., 1999) it is clear that these additional domains have an as yet unexplained inhibitory role with respect to Chd. In an experiment localising the injection of *Xld* and *XldC1-3* to the ventral marginal zone, the latter appeared to have a greater ventralising ability than *Xld* (Figure 4.5, compare B & C). This might suggest that last two CUB domains (CUB4 and CUB5) and/or the last EGF domain (E2) have an inhibitory role in Xld with respect to its chordinase activity. Alternatively, and a more trivial explanation, is that *XldC1-3* was translated more efficiently than *Xld*, which may explain this result (see Figure 4.3, lanes B & G). To test this, the amount of mRNA injected into embryos was optimised so that *Xld* and *XldC1-3* produced similar amounts of protein, as determined on Western blots. Initially 0.8, 1.6, and 3.2ng of *Xld* mRNA (Figure 4.6B-D) and 0.025, 0.05, and 0.1ng of *XldC1-3* mRNA (Figure 4.6E-G) was injected into *Xenopus* embryos at the 2-cell stage. Embryos were cultured until stage 10/10.5 and five embryos removed for Western blot analysis (Figure 4.6Ai). The remaining embryos were cultured until stage 29-30 to determine the extent of ventralisation indicated by DAI score. Injection of 3.2ng *Xld* mRNA (Figure 4.6D) gave a similar ventralised phenotype to that of 0.025ng *XldC1-3* mRNA (Figure 4.6E) with

Figure 4.5 Ventral injection of *Xld* and *XldC1-3*

0.8ng mRNA was injected into two ventral blastomeres at the 4 to 8 cell stage for either *Xld* (B) or *XldC1-3* (C) and cultured until uninjected sibling controls had reached stage 26/27 (A). Ventral injection of *XldC1-3* appeared to give a more ventralised phenotype (33/33 cases) when compared to ventral injection of *Xld* (8/37 cases) (compare B & C).

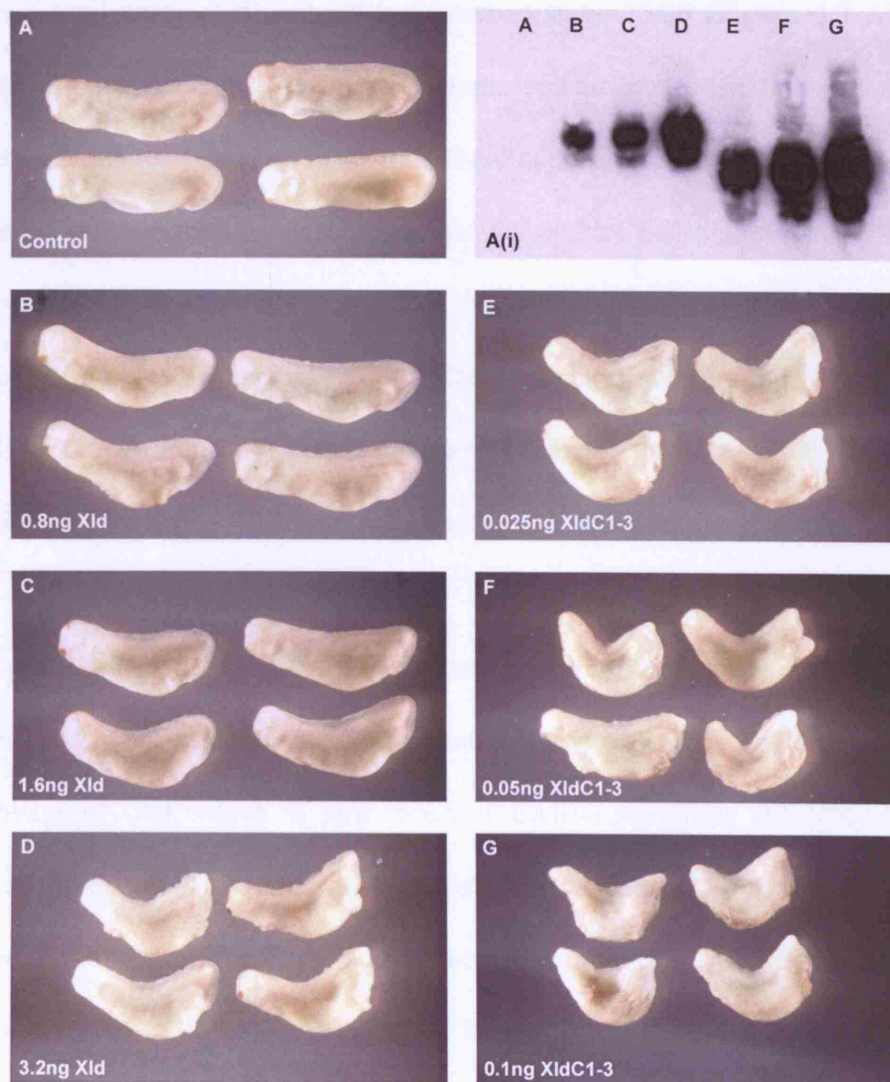
Figure 4.6 Optimisation of *Xld* and *XldC1-3* injections

Embryos were injected at the two cell stage with 0.8ng (B), 1.6ng (C), and 3.2ng (D) *Xld* mRNA or 0.025ng (E), 0.05ng (F), and 0.1ng (G) *XldC1-3* mRNA. Embryos were cultured until uninjected sibling embryos had reached stage 27/28. At stage 10/10.5 five embryos were taken and subjected to Western blot analysis to determine the level of translation (A(i)). 3.2ng of *Xld* appears to give the same phenotype at 0.025ng of *XldC1-3* (compare D & E). A calculation of the DAI for these two sets of injections showed a score of 3.9 for *Xld* and 3.4 for *XldC1-3* and Western blot analysis shows that these two concentrations give approximately the same amount of translated protein (compare A(i) lanes D & E).

Figure 4.5



Figure 4.6



DAIs of 3.9 and 3.4 respectively. Comparison of protein levels by Western blot showed that these amounts of mRNA gave similar levels of protein (Figure 4.6A(i), lanes D & E). This result indicates that it is a difference in translational efficiency, and not activity, that is responsible for increased ventralisation by *XldC1-3*, when compared to *Xld*, in Figure 4.5. To test this further, either 3.2ng of *Xld* mRNA or 0.025ng of *XldC1-3* mRNA was injected into either two dorsal (Figure 4.7C, E) or two ventral (Figure 4.7D, F) blastomeres at the 4- to 8-cell stage and the embryos grown until uninjected controls had reached stage 25/26 (Figure 4.7 A). Again, a small number of embryos were homogenised at stage 10/10.5 and subjected to Western blot analysis (Figure 4.7B). There appeared to be no difference between the level of ventralisation produced by either *Xld* or *XldC1-3* regardless of the location of injection. The Western blot confirmed that the levels of protein produced for each set of injections was similar and indicates that removal of the 4th and 5th CUB domains, as well as the 2nd EGF-like domain, does not enhance the activity of *Xld* and that the initial result was due to differential translation of the mRNAs.

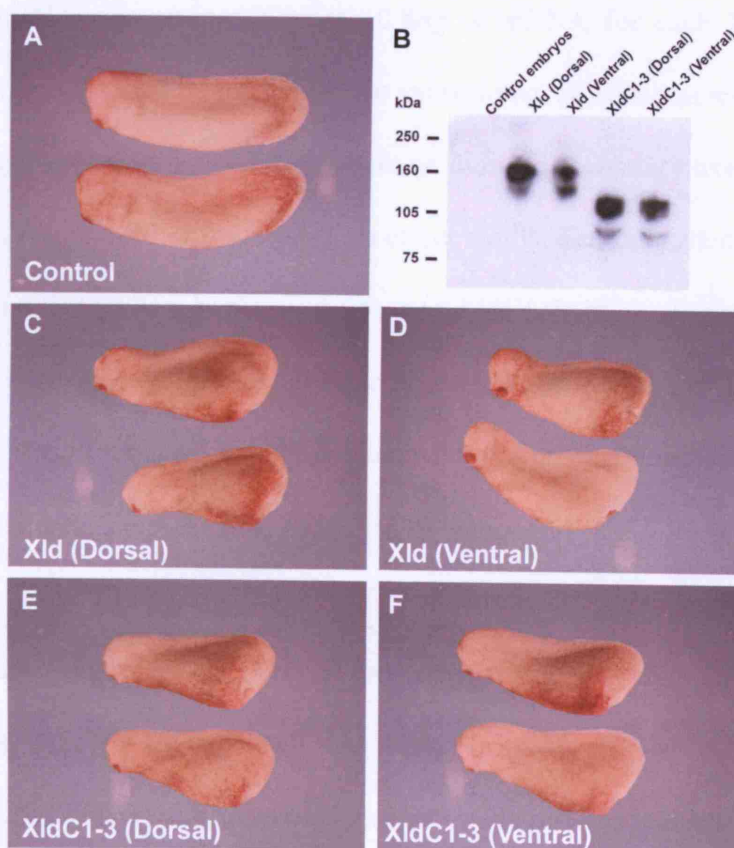
4.2.3 Both CUB1 and CUB2 of *Xld* are required to counteract co-injected *Chd*

The most likely explanation for the ventralisation caused by *Xld* is due to increased cleavage of *Chd*, which in turn enhances BMP-4 signalling (Piccolo et al., 1997). Over-expression of *chd* mRNA results in the expansion of dorsal cell identities, at the expense of ventral identities, due to increased inhibition of BMP-4 in the marginal zone (Piccolo et al., 1996; Sasai et al., 1994). When *Chd* is ectopically expressed in the VMZ of *Xenopus* embryos it can mimic aspects of Spemann and Mangold's organiser transplant

Figure 4.7 Dorsal and Ventral injection of *Xld* and *XldC1-3*

Embryos were injected with either 3.2ng *Xld* mRNA or 0.025ng *XldC1-3* mRNA into two dorsal blastomeres (C & E) or two ventral blastomeres (D & F) at the 4 to 8 cell stage. At stage 10/10.5 five embryos were taken and 1 embryo worth of protein extract subjected to Western blot analysis (B). Remaining embryos were cultured until uninjected sibling controls had reached stage 25/26 (A). Each set of injections gave an identical phenotype, regardless of the injection site (n=45 (C), n=42 (E), n=51 (E), n=43 (F)). The Western blot confirmed that similar levels of protein were produced for each mRNA.

Figure 4.7

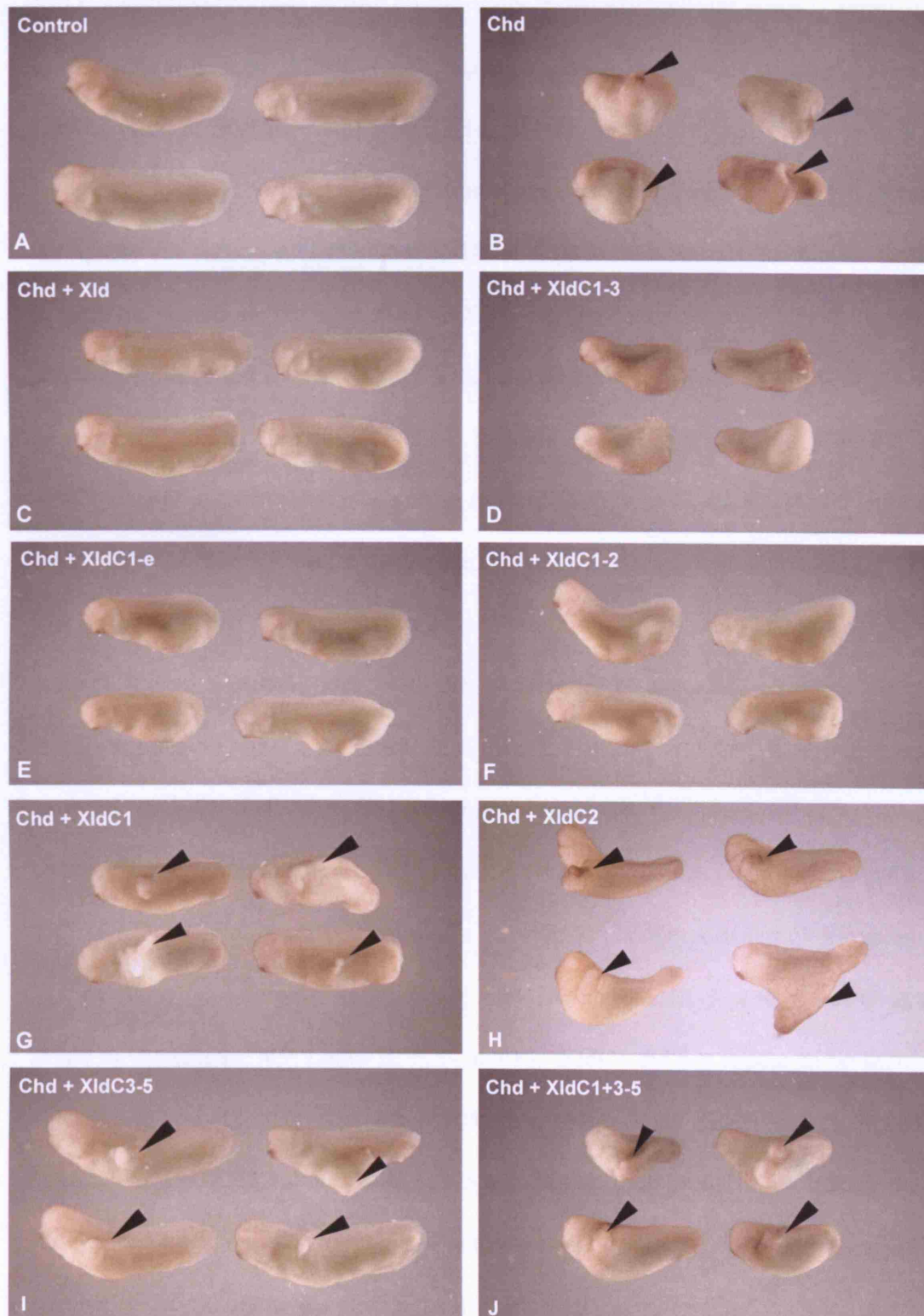


and induce a secondary dorsal axis, albeit without a complete head (Sasai et al., 1994). These secondary axes can be counteracted by co-injection either of *xbmp-1*, *Xld* or *xlr* mRNA (Dale et al., 2002; Goodman et al., 1998; Piccolo et al., 1997; Wardle et al., 1999). To confirm that the ventralised phenotype obtained by *Xld*, *XldC1-3*, *XldC1-e* and *XldC1-2* injection was due to increased cleavage of Chd, 0.8ng of *chd* mRNA was co-injected with 0.8ng of mRNA for each *Xld* CUB deletion construct. mRNA was injected into two ventral-vegetal blastomeres at the 4- to 8-cell stage. Injection of 0.8ng *chd* mRNA alone induced secondary axes in 58% of cases (80/139) (Figure 4.8B, arrowheads), an effect that was counteracted by coinjected *Xld* (0/77), *XldC1-3* (0/74), *XldC1-e* (0/65), and *XldC1-2* (0/74) (Figure 4.8 C-F; Figure 4.9). In most cases, embryos co-injected with *Xld* mRNAs exhibited a ventralised phenotype. *Chd* induced secondary axes were not counteracted by coinjected *XldC1* (45% secondary axes, 37/82), *XldC2* (69% secondary axes, 24/35), *XldC3-5* (28% secondary axes, 21/76), or *XldC1+3-5* (76% secondary axes, 56/74) (Figure 4.8G-I, Arrowheads; Figure 4.9). This is consistent with the absence of a ventralised phenotype in embryos injected with these mRNAs (Figure 4.2). It was noted, however, that *XldC3-5* still appeared to retain some residual activity against Chd, with only 28% induction of secondary axes. These results confirm that both the 1st and 2nd CUB domains are required for *Xld* activity against Chd.

Figure 4.8 Coinjection of *Xld* deletion constructs and *chd* mRNA

Ventral injection of 0.8ng *chd* mRNA at the 4 to 8 cell stage results in the formation of secondary axes in 58% (80/139) of embryos (B). These secondary axes were counteracted by co-injection of either, *Xld* (0/77 embryos) (C), *XldC1-3* (0/74 embryos) (D), *XldC1-e* (0/65 embryos) (E), or *XldC1-2* (0/74 embryos) (F) mRNA. Secondary axes were present when *chd* was co-injected with either *XldC1* (37/82 embryos) (G), *XldC2* (26/35 embryos) (H), *XldC3-5* (21/76 embryos) (I), or *XldC1+3-5* (56/74 embryos) (J) mRNA. Secondary axes indicated by black arrowheads.

Figure 4.8



4.2.4 The presence of CUB2 is required for secretion of Xld

Xld is a secreted protease, acting on substrates in the extracellular space (Goodman et al., 1998; Wardle et al., 1999). One explanation for the loss of ventralising activity by XldC1, XldC2, XldC3-5, and XldC1+3-5, therefore, is that these proteins are not secreted. Consistent with this explanation is the observation that the 1st CUB domain is required for secretion of mammalian BMP-1, showing a role for the CUB domains in protease secretion (Hartigan et al., 2003). To determine whether any of the CUB domains are required for secretion of Xld, each deletion construct, and full-length *Xld*, was transfected into HEK293T cells. In addition, new versions of XldC1-2 and XldC1-e were produced that included an in frame Myc tag sequence. Media and cells were collected after 48 hours incubation, proteins extracted and separated by SDS-PAGE. Myc-tagged Xld was detected by Western blot using the 9E10 antibody. All constructs were translated well, with the Myc-tagged proteins being readily detected in cell extracts (Figure 4.9, Lanes A, C, E, G, I, K, M and O). The major band detected in each cell extract lane corresponded to the predicted size of the unprocessed deletion construct, including the Myc tag sequence (summarised in Table 4.1). Of those deletions that displayed a ventralised phenotype when over expressed in the embryo, full-length Xld, XldC1-2, XldC1-e, and XldC1-3 protein was detected in the media (Figure 4.9, Lanes B, F, N and P). XldC1 did not appear to be secreted at all (Lane D) but low levels were detected for XldC3-5 and XldC1+3-5 (Lanes H & J). Secretion of XldC3-5 and XldC1+3-5 was similar to that of Xld, indicating that it should be sufficient for biological activity. In contrast, XldC2 was secreted efficiently yet did not give a phenotype in the

Figure 4.9 Secretion of Xld deletions

2.5 µg of each deletion construct was transformed into HEK293T cells cultured for 24 hours in DMEM with serum then 24 hours in serum free DMEM. The serum free media was collected and the cells homogenised in 2ml of homogenisation buffer. 2µl of cell lysate and 10µl of serum free media was subjected to Western blot analysis to determine which deletion constructs were secreted into the media. All constructs were translated well by the cells (lanes A, C, E, G, I, K, M, O). Protein for Xld, XldC1-3, XldC2, XldC1-2 and XldC1-e was detected in the media (lanes B, F, L, N, P). A small single band of protein was detected in the media for XldC3-5 and XldC1+3-5 (lanes H and J) but this is likely to be overspill from adjacent lanes. No protein was detected in the media for XldC1 (lane D).

Table 4.1 Predicted molecular weights of Xld CUB deletions

The predicted molecular weight of each Xld CUB deletion is given in kDa. The unprocessed form includes both the signal sequence and pro-peptide which comprises approximately 20 kDa. The processed column gives the sizes of each Xld deletion if the pro-peptide and signal sequence are removed. Proteins of this size should correspond to an active protease.

Figure 4.9

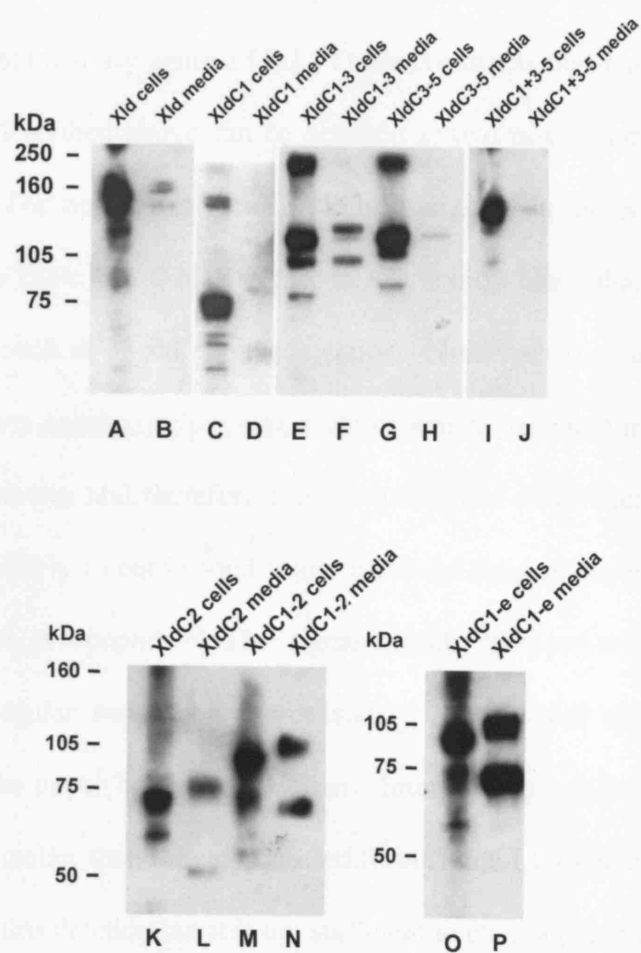


Table 4.1

	Unprocessed	Processed
Xld.MT	123	103
XldC1/C2	71	50
XldC1-2	83	63
XldC1-e	88	68
XldC1-3	101	81
XldC3-5	101	81
XldC1+3-5	114	94

whole embryo (Lane L), demonstrating that the second CUB domain is required for secretion but not activity against Chd. Those constructs that have biological activity in the assays described above can be detected as two bands in culture media (Lanes, B, F, N, P). The uppermost band is slightly larger than the predicted protein size, summarised in Table 4.1. This increase in size is most likely due to post-translational modifications such as N-linked glycosylation. Nine such sites are predicted for Xld by Prosite (www.expasy.org/prosite), eight of which are found in the first 400 amino acids of the protein and therefore present in each of these secreted products. The lower band is likely to correspond to the activated form of the protease, which lacks the N-terminal pro-peptide. The signal-peptide and pro-region of Xld have a predicted molecular weight of approximately 20 kDa and these lower bands are smaller than the upper band by that order. Interestingly, media for XldC2 reveals a band 20 kDa smaller than the unprocessed form (Lane L), indicating that the protease is activated in this deletion but it is not sufficient to elicit a phenotype in the embryo.

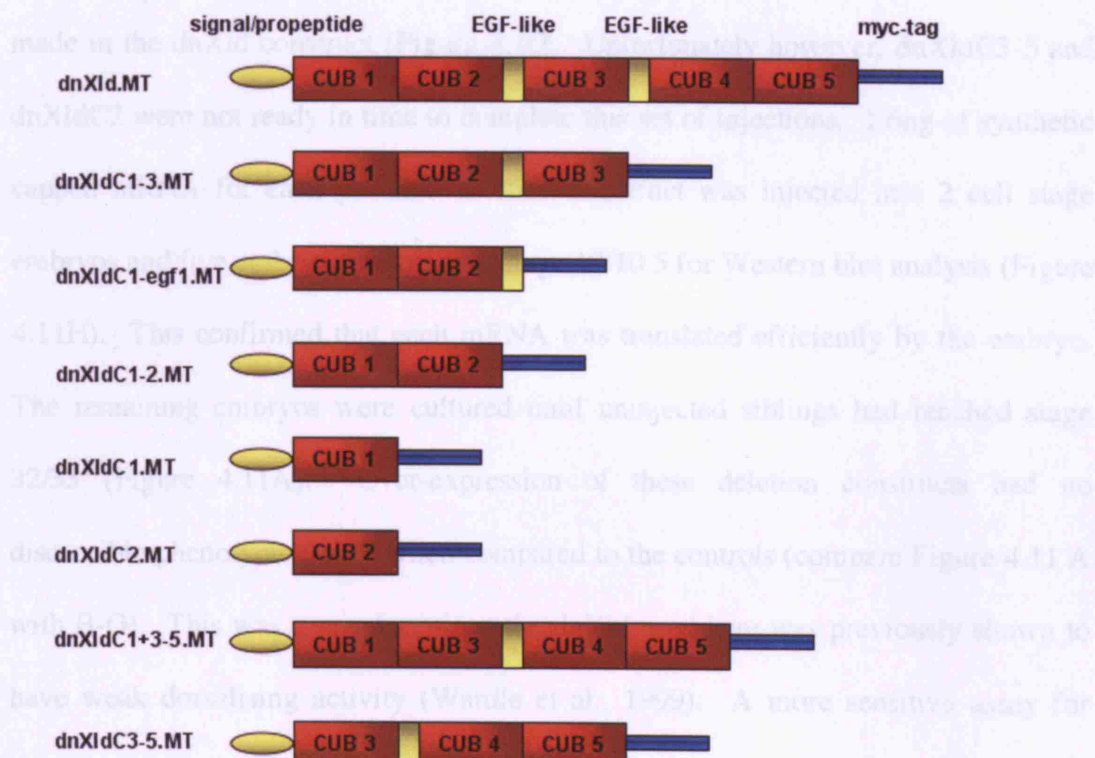
4.2.5 Dominant negative Xld constructs

Removal of the N-terminal protease domain from XBMP-1 and Xld, or its inactivation by point mutation, results in the formation of a dominant-negative version of the protease (Piccolo et al., 1997; Wardle et al., 1999). These dominant-negative proteases are able to dorsalise ventral mesoderm when injected into early *Xenopus* embryos. If only the first and second CUB domains are required for Xld function against Chd, then one would expect that only those two domains would be required to dorsalise ventral mesoderm in the dominant-negative version. In an attempt to determine if this was the case, I utilised

Figure 4.10 Schematic diagram of putative dnXld CUB deletions

The protease domain was removed from the Xld CUB deletion constructs described in Figure 4.1, in an attempt to produce dominant-negative forms for each. Again, each construct included the signal/propeptide and 6 Myc tag epitopes at the C terminus.

Figure 4.10



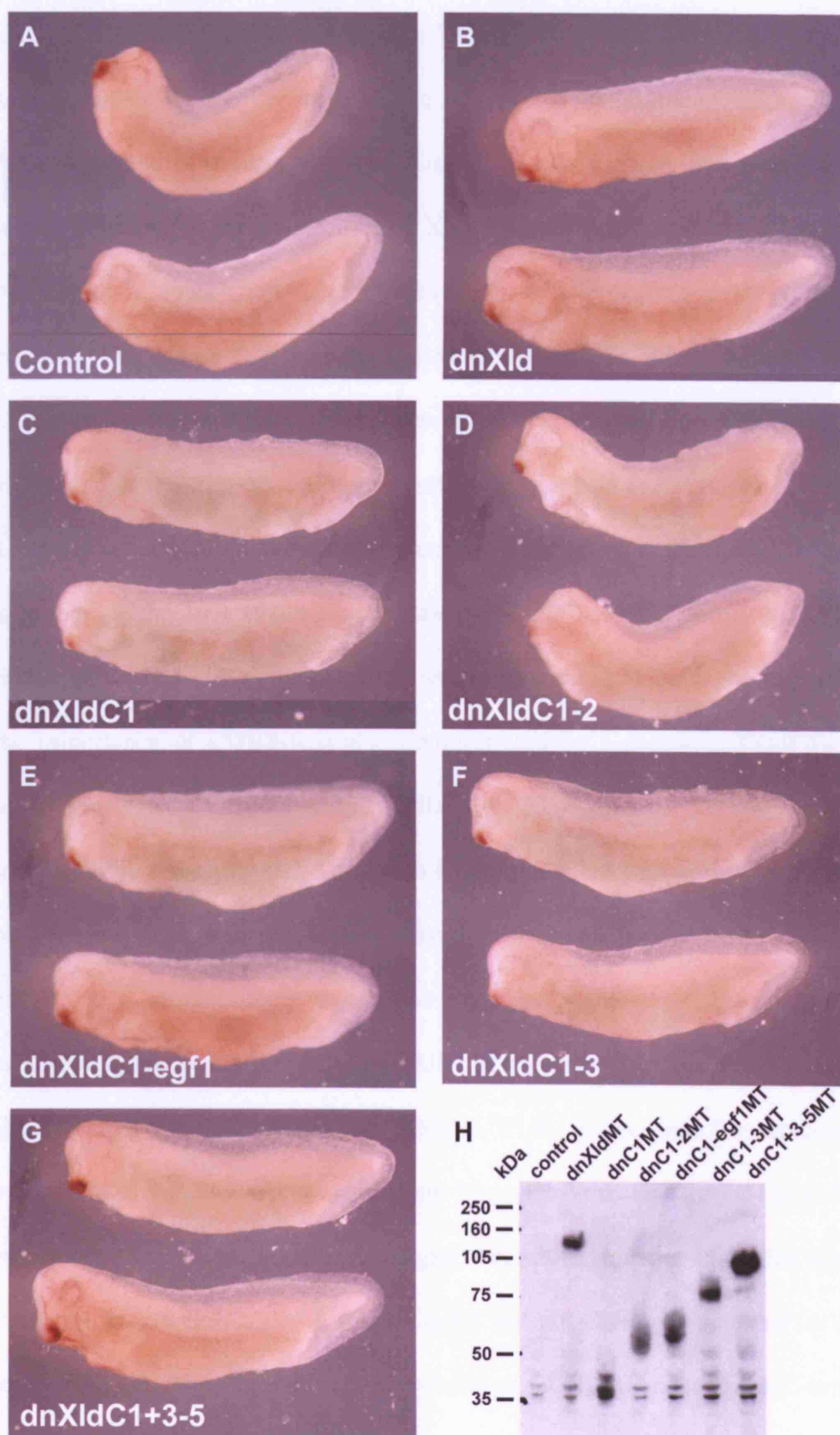
the dnXld construct built by Fiona Wardle (Wardle et al., 1999). This construct combines the C-terminal CUB domains of Xld with the N-terminal signal sequence of BMP-1. The pro-peptide was also eliminated since it contains an unpaired cysteine residue whose effect is at present unknown. The same PCR approach as used to make the original CUB deletions (Figure 4.1) was employed, and a set of identical deletions made in the dnXld construct (Figure 4.10). Unfortunately however, dnXldC3-5 and dnXldC2 were not ready in time to complete this set of injections. 1.6ng of synthetic capped mRNA for each protease deletion construct was injected into 2 cell stage embryos and five embryos collected at stage 10/10.5 for Western blot analysis (Figure 4.11H). This confirmed that each mRNA was translated efficiently by the embryo. The remaining embryos were cultured until uninjected siblings had reached stage 32/33 (Figure 4.11A). Over-expression of these deletion constructs had no discernible phenotypic defect when compared to the controls (compare Figure 4.11 A with B-G). This was a surprise, since the dnXld used here was previously shown to have weak dorsalising activity (Wardle et al., 1999). A more sensitive assay for dorsalisation is to isolate VMZs from injected embryos and Wardle et al. (1999) have shown that dnXld strongly dorsalises at least 1/3rd of VMZs in this assay. Consequently, embryos were injected with 1.6ng of mRNA, for each dominant-negative deletion construct, into 2 ventral-vegetal blastomeres at the 4- to 8-cell stage. VMZs were isolated at stage 10/10.5 and cultured until sibling embryos reached stage 31/32. Total RNA was extracted and analysed by both RT-PCR and RNase protection for the presence of the dorsal muscle marker *α -actin*, which would indicate dorsalisation of the VMZ. In each assay the injected VMZs were indistinguishable from control VMZs (data not shown). In a further attempt to determine if dorsalisation had occurred, VMZ explants injected with each dominant-negative were

cultured until stage 40 and histologically sectioned to identify muscle or notochord structures. Although muscle appeared to be present in *dnXldC1-e* injected explants, none was present in those injected with *dnXld*.

Figure 4.11 Injection of mRNA for dnCUB deletions

0.8ng of synthetic capped mRNA, for either *dnXld* (B), *dnXldC1* (C), *dnXldC1-2* (D), *dnXldC1-e* (E), *dnXldC1-3* (F), or *dnC1+3-5* (G), was injected into each blastomeres at the two-cell stage. Embryos were grown until uninjected siblings reached stage 31/32 (A). At stage 10/10.5, 5 five embryos were homogenised and 1 embryos worth of protein extract subjected to Western blot analysis to check for translation (H). All constructs translated efficiently but gave no phenotype when injected into the whole embryo (n=16 (B), n=28 (C), n=15 (D), n=17 (E), n=24 (F), n=20 (G)). At the time of conducting this experiment *dnXldC3-5* and *dnXldC2* were not yet made.

Figure 4.11



4.3 Discussion

4.3.1 CUB and EGF-like domains in Xld

In this chapter I have investigated the role of the C-terminal CUB and EGF-like domains in controlling the action of Xld. PCR was used to introduce a number of module deletions in the C-terminus of Xld, a region that is thought to play a role in protein-protein interactions and consequently convey substrate specificity to the protease. Full-length Xld contains five CUB and two EGF-like domains and, when ectopically expressed in *Xenopus* embryos, causes mild ventralisation of the mesoderm resulting in an enlarged posterior region and reduced head (Goodman et al., 1998). A similar phenotype was observed for deletion constructs that contained a minimum of the first two CUB domains (XldC1-2), while the first (XldC1) and second (XldC2) CUB domains alone were insufficient to generate this phenotype. The importance of CUB2 was also demonstrated by a construct (XldC1+3-5) that lacked only this CUB domain, and the first EGF domain (EGF1), and did not ventralise injected embryos. Thus both CUB1 and CUB2 appear to be necessary for the ventralizing activity of Xld. A trivial explanation for the failure of some Xld mutants to ventralise injected *Xenopus* embryos is that they are not secreted, as suggested by the demonstration that CUB1 is necessary for secretion of mammalian BMP-1 (Hartigan et al., 2003). In BMP-1, CUB1 is required for secretion of the protease from EBNA293 cells and must be in its correct position, immediately C-terminal of the protease domain (Hartigan et al., 2003). Using a similar approach in this study, I have shown that for Xld CUB2, and not CUB1, is required for secretion from HEK-293 cells. In addition, movement of CUB2 adjacent to the C-terminus of the protease domain still allows secretion, indicating its position is not as important as that of CUB1 in BMP-1.

Since Xld is known to cleave Chd (Piccolo et al., 1997; Wardle et al., 1999), the most likely explanation for the results described in this chapter is that CUB1 and CUB2 are necessary for cleavage of this dorsalizing factor. To test this, each deletion construct was co-injected into ventral mesoderm with Chd, which induces a second dorsal axis, albeit lacking head structures, when injected alone (Sasai et al., 1994). In this assay only Xld CUB deletions that ventralised embryos were able to block the dorsalizing activity of Chd. This demonstrates that both CUB1 and CUB2 are both necessary and sufficient for Xld to block the activity of Chd, indicating that these domains are responsible for targeting Xld to this important substrate. It should be possible to demonstrate this directly by co-precipitating versions of XldC1-2 that lack protease activity with Chd, as was recently demonstrated with Tlr/Tlk and Sog in *Drosophila* (Serpe et al., 2005). It would also be advisable to show that XldC1-2 cleaves Chd *in vitro*, by adding XldC1-2 conditioned media to purified Chd protein. Unfortunately time constraints have not permitted these two experiments.

Genetic studies in *Drosophila* suggest that the second CUB domain is also necessary for the activity of Tld in dorsal-ventral patterning at the cellular blastoderm stage (Childs and O'Connor, 1994; Finelli et al., 1994). By constructing *tld* and *dpp* double mutants, these authors identified a small number of *tld* mutations that enhanced the severity of the *dpp* phenotype. These mutations were mapped to either the protease domain or CUB2 domain, while mutations in CUB3, CUB4, CUB5, and EGF2 had no effect on the *dpp* phenotype. This indicated that the protease and CUB2 domain modulated Dpp activity and that the remaining C-terminal domains are less important. At the time of these studies a molecular mechanism for the ability of some *tld*

mutations to modulate *dpp* was not clear, but it is now apparent that the activity is mediated by the Dpp binding protein, and inhibitor, Sog (Marques et al., 1997). Consequently we can see that *Drosophila* Tld requires both the protease domain and CUB2 for this anti-Sog function.

Recent biochemical studies of mammalian BMP-1 have also shown CUB2 to be important for the processing of other Tld substrates. Hartigan et al. (2003) identified BMP-1 CUB2 as being required for the C-terminal processing of proColl α 1 and Petropoulou et al. (2005) showed that the same domain was required for Chordinase activity *in vitro*. These studies suggest the function of CUB2 is conserved between different Tld family members and across different species. In addition to demonstrating that CUB2, and not CUB1, is required for Xld secretion, I have shown that in order for Xld to act against Chd, CUB2 must be separated from the protease domain by the presence of an intervening CUB domain. From these experiments we cannot say if CUB1 specifically is required or if any CUB domain will suffice, but clearly the presence of CUB2 in its correct position is important. While beyond the scope of this study, it would be interesting to determine if CUB2 of Xld could compensate for the lack of chordinase activity displayed by other members of the Tld family. Mammalian Tll-2 does not cleave Chd (Scott et al., 1999), however it can be converted to a chordinase by swapping its entire CUB array for the CUB1 domain of BMP-1 (Petropoulou et al., 2005). Since CUB1 seems to be critical for the activity of BMP-1, would CUB2 of Xld, critical for the latter's activity, produce a Chd cleaving Tll-2 in a similar experiment?

The presence of two calcium binding EGF-like domains in Xld, as in the other Tlds, suggests that the binding of calcium may play a role in regulating protease activity or CUB domain interactions. However, this study and those involving mammalian BMP-1 (Hartigan et al., 2003; Petropoulou et al., 2005) indicate that EGF-like calcium binding is dispensable for the purposes of Chd and procollagen cleavage. Whether their presence is required for the processing of other substrates has yet to be investigated. Intriguingly, removal of both EGF-like domains from mTld allows the protease to cleave Chd, which the wildtype is unable to do, in the absence of Tsg, and enhances its procollagen processing activity (Garrigue-Antar et al., 2004). This suggests that the presence of two EGF-like domains play an inhibitory role in mTld but this is not the case for Xld or other Tld family members with a similar domain structure.

4.3.3 Dominant-negative constructs

Previous experiments have shown that removal of the protease domain from Xld, or its inactivation, creates a dominant-negative protease capable of dorsalising ventral mesoderm or the embryo as a whole (Blitz et al., 2000; Piccolo et al., 1997; Wardle et al., 1999). To determine the modules responsible for this dominant-negative activity the protease domain was removed from each of the deletion constructs. While all were translated well *in vivo*, none produced a phenotype when expressed in *Xenopus* embryos and did not dorsalize isolated VMZs. The reasons for the difference between these results, and those obtained by previous authors are unknown but may relate to the presence of the tag employed here. All the constructs used here had 6 Myc tags added to the C-terminus, while previous studies used constructs with only a single C-terminal Flag tag (Piccolo et al., 1997) or no tag at all (Wardle et al., 1999).

While the tags do not affect the protease activity of either wildtype Xld or the Xld CUB deletions, we now suspect that they are interfering with the previously reported dominant-negative activity. Untagged versions of the deletion constructs have now been made but were not ready in time to complete this study.

4.3.4 Summary

In this chapter I have demonstrated that both CUB1 and CUB2 of Xld are required for both secretion and anti-Chd activity of the protease. Alone, neither domain is sufficient. I have also shown that the EGF-like domains are dispensable for Xld action against Chd, but we do not know if they are required to cleave other substrates. In the future I hope to show that deletion constructs, lacking a protease domain, can act as dominant negatives for their corresponding active enzymes and perform co-precipitation experiments to demonstrate which domains physically interact with Chd. A table summarising these results can be found in Figure 4.12.

Figure 4.12 Summary of results obtained with each CUB deletion construct.

A table summarising the results obtained with each deletion construct. The minimal domain structure to replicate all the results obtained with full length Xld is XldC1-2 comprising of the protease domain and the first and second CUB domains. XldC2 comprising of the protease and only the second CUB domain is secreted and processed to a mature form but is unable to act against Chd in these assays. A similar construct containing the protease and only CUB1 is unable to reproduce any of the activity we have detected for Xld or XldC1-2.

Figure 4.12

Endoderm is not an inhibitor of Tld gradient.

5.5 Introduction

5.5.1 The Tld during dorsal-ventral patterning

In *Xenopus*, dorsal-ventral patterning of the embryo occurs, in part, due to inhibition of the ventralizing signaling pathway by factors secreted from the organizer (Oziri and Kessler, 1997). For normal development to proceed there must be a balance of BMP-4 and Dpp. The organizer, otherwise the embryo will be ventralized (Kessler and Kessler, 1992). Factors secreted by the organizer include chordin (Chd), noggin, and follistatin, all of which directly bind BMP-4 and prevent it from activating BMP receptors (Famoud et al., 1997; Piccolo et al., 1996; Zimmerman et al., 1996). Injection of mRNA for either of these molecules dorsalises embryo and ectopic expression in the ventral region induces secondary axes, mimicking the effect of the organizer (Hemmati-Brivanlou et al., 1994; Sasai et al., 1994; Smith and Melnick, 1992). Chd is of particular interest since it is evolutionary conserved in *Drosophila* and is required for dorsal-ventral patterning in the blastoderm stage embryo (Zimmerman et al., 1998). In *Drosophila*, the Chd ortholog secret gastrulation directly binds and inhibits Decapentagene (Dpp), itself an ortholog of ventral BMP-4 that is produced by the dorsal ectoderm of the

Drosophila embryo (Hargrove and Anderson, 1992). Chd and Scg are both regulated by members of the Tld family, secreted inhibitors, which change their inhibitors releasing fragments with reduced affinity for BMP-4 and Dpp (Carram et al., 2000; Marquis et al., 1997; Piccolo et al., 1996). In this way BMP-4 and Dpp become free to activate their receptors. Alternatively, transcripts for two members of the Tld family, Abop-1 and Tld, are detected in the organizer of *Chironomus tentans*

(Carram et al., 2000). Chd is of particular interest since it is evolutionary conserved in *Drosophila* and is required for dorsal-ventral patterning in the blastoderm stage embryo (Zimmerman et al., 1998). In *Drosophila*, the Chd ortholog secret gastrulation directly binds and inhibits Decapentagene (Dpp), itself an ortholog of ventral BMP-4 that is produced by the dorsal ectoderm of the

Drosophila embryo (Hargrove and Anderson, 1992). Chd and Scg are both regulated by members of the Tld family, secreted inhibitors, which change their inhibitors releasing fragments with reduced affinity for BMP-4 and Dpp (Carram et al., 2000; Marquis et al., 1997; Piccolo et al., 1996). In this way BMP-4 and Dpp become free to activate their receptors. Alternatively, transcripts for two members of the Tld family, Abop-1 and Tld, are detected in the organizer of *Chironomus tentans*

Chapter 5

Endodermin is not an inhibitor of Tld proteases

5.1 Introduction

5.1.1 The Tlds during *Xenopus* gastrulation

In *Xenopus*, dorsal-ventral patterning of the marginal zone occurs, in part, due to inhibition of the ventralising signalling molecule BMP-4 by factors secreted from the organiser (Dale and Jones, 1999; Dale and Wardle, 1999). For normal development to proceed there must be maximal inhibition of BMP-4 in the organiser, otherwise the embryo will be ventralised (Dale et al., 1992; Jones et al., 1992). Factors secreted by the organiser include chordin (Chd), noggin, and follistatin, all of which directly bind BMP-4 and prevent it from activating BMP receptors (Fainsod et al., 1997; Piccolo et al., 1996; Zimmerman et al., 1996). Injection of mRNA for either of these molecules dorsalises embryos and ectopic expression in the VMZ induces secondary axes, mimicking the effect of the organiser (Hemmati-Brivanlou et al., 1994; Sasai et al., 1994; Smith and Harland, 1992). Chd is of particular interest since it is evolutionary conserved in *Drosophila* and is required for dorsal-ventral patterning in the blastoderm stage embryo (Zusman et al., 1988). In *Drosophila*, the Chd orthologue short gastrulation (Sog) directly binds and inhibits decapentaplegic (Dpp), itself an orthologue of vertebrate BMP-4 that is required to pattern the dorsal ectoderm of the *Drosophila* embryo (Ferguson and Anderson, 1992a). Chd and Sog are both regulated by members of the Tld family of metalloproteases, which cleave these inhibitors releasing fragments with reduced affinity for BMP-4 and Dpp (Larrain et al., 2000; Marques et al., 1997; Piccolo et al., 1997). In this way BMP-4 and Dpp become free to activate their receptors. Paradoxically, transcripts for two members of the Tld family, *Xbmp-1* and *Xld*, are detected in the organiser of *Xenopus* gastrulae

(Goodman et al., 1998) and if maximal inhibition of BMP-4 by Chd is required in the DMZ then the presence of these proteases maybe detrimental to that process. In thinking about a resolution to this problem we envisaged a scenario whereby an inhibitor of Tld metalloproteases was present in the organiser, to protect Chd from unnecessary cleavage. One candidate for such a function that came to our attention was Endodermin (Edd).

5.1.2 Endodermin and α -2 Macroglobulin

Edd was first identified in a screen for genes specifically expressed in the organiser region of *Xenopus* gastrulae (Sasai et al., 1996). Transcripts are found throughout the marginal zone of early gastrulae and enriched in the organiser. After gastrulation, levels of *edd* expression increase and transcripts are found throughout the endoderm, it is also expressed in the mesodermally derived notochord. Edd is a 1461 amino acid secreted protein most closely related to α -2 macroglobulin (α -2M), the most abundant protein in human blood. Comparison of amino acid identities shows high homology between the two proteins; overall Edd is 49% identical to α -2M, though some regions show as much as 63% identity. α -2M is a non-specific protease inhibitor that binds enzymes using a trapping mechanism. It contains a 40 amino acid “bait” region that is highly susceptible to proteolytic cleavage and contains multiple cleavage sites for serine, cysteine, and aspartic proteases, as well as for metalloproteases such as collagenases (Barrett, 1981; Sottrup-Jensen et al., 1989). Upon bait region cleavage a downstream thio ester group becomes highly reactive and covalently binds the cleaving protease, in turn the two halves of α -2M surround and trap the enzyme preventing further substrate access. The protease however remains catalytically active and small substrates capable of penetrating the complex can be cleaved (Sottrup-

Jensen et al., 1989). Edd contains a 47 amino acid sequence whose position corresponds to the bait region of α -2M, but homology between the two regions is low – they are only 15% identical. The bait region of α -2M does include a single Asp residue, which is conserved in Edd (Figure 5.1) and residues in the surrounding region that are often found in the cleavage sites of Tld substrates (see Chapter 1). As yet there are no reports of a receptor for Edd, or indeed any other proteins that Edd might complex with.

5.1.3 Proteases inhibited by α -2M

α -2M inhibits a wide range of proteases, including many belonging to the metzincin superfamily of proteases. The metzincins, characterised by their distinctive “Met turn”, includes the matrix metalloproteases (MMPs), A Disintegrin And Metalloproteinases (ADAMs) and the Tolloids (Tld). α -2M binds many members of this superfamily including MMP-1 (also known as collagenase 1), MMP-9 (or Type IV collagenase) and ADAMTS 4 and 5 (Grinnell et al., 1998; Ogata et al., 1992; Tortorella et al., 2004). While there is no direct evidence that α -2M inhibits members of the Tld family, an α -2M like protein isolated from the haemolymph of the Crayfish *Astacus astacus* was found to inhibit Astacin (Stocker et al., 1991). Astacin is very closely related to the protease domain of the Tlds and thought to be the prototype enzyme of this group, raising the possibility that an α -2M like protein does indeed inhibit these proteases.

Figure 5.1 A comparison of the bait region of Edd and α -2M shows little conservation

The 36 amino acid putative bait region of Edd is only 15% similar to the 39 amino acid bait region of human α -2M. Amino acids highlighted in red are identical, or a conservative change, in both proteins. There is a conserved aspartic acid (D) residue (blue box) in both proteins, an amino acid that is always found in the P1 position of all known BMP-1 cleavage sites.

Figure 5.1

15%

Edd	IRIPVVC	EERLG---	RGAVQFSSFA	VSADFG	EMAFMALMHSTEVI	ETVRKY
α-2M	IRKPKK	PQLQQY	EMHGPEGLRVGF	YESDVG	RGHARLVHVEEPHT	ETVRKY

5.1.4 Aims

Edd is a *Xenopus* protein closely related to α -2M, a non-specific protease inhibitor that attenuates the proteolytic activity of a number of members of the metzincin protease superfamily. The Tld related metalloproteases also belong to this family and during *Xenopus* gastrulation are expressed in the same region as *edd*. In this chapter I investigate the possibility of Edd acting as an inhibitor of the Tld protease XBMP-1. I also determine whether α -2M itself inhibits Xlr, a Tld protease closely related to XBMP-1.

5.2 Results

5.2.1 Expression of *edd* by whole mount *in situ* hybridisation

To investigate the role that *edd* plays in *Xenopus* development we first confirmed the expression pattern of transcripts for the gene. Sasai et al. (1996) reported that *edd* is first detected at gastrulation, in the marginal zone with enrichment of transcripts in the dorsal sector (DMZ). Later expression is then found throughout the endoderm and notochord. Our experiments confirmed this; during gastrulation *edd* is expressed throughout the marginal zone and staining for transcripts is indeed enriched in the DMZ (Figure 5.2A). Transcripts appear localised to the nuclei of cells, some of which are thought to be fated to become endoderm following gastrulation (Sasai et al., 1996). After gastrulation *edd* expression increases and becomes localised to the endoderm with a low level of expression found in the notochord. It is however completely excluded from the neural plate, neural tube and somites (Figure 5.2B-D). A closer inspection of the expression pattern of *edd* at gastrulation was carried out and compared to that of *chd* (Figure 5.3). *Chd* expression is restricted to the DMZ (Figure 5.3A), with heavy staining indicating high transcript levels. *edd* expression is less intense than *chd* and could be described as speckled, transcripts appearing in distinct cells peppered throughout the marginal zone (Figure 5.3B). Even where transcripts are enriched in the DMZ, expression is seen in distinct cells unlike *chd* (compare Figure 5.3A & B). An unexpected result arose with observation of *edd* expression in a small number of cells that appeared to be outside the marginal zone (Figure 5.3B). These cells appear to be further away from the main expression domain of *edd* than we expected and might even be outside what is traditionally thought to be the marginal zone.

Figure 5.2 Spatial expression of *edd* during embryogenesis

During gastrulation, *edd* transcripts are found throughout the marginal zone in a speckled fashion, with enrichment in the DMZ (A). At stage 18/19 expression is found throughout the endoderm and is excluded from the neural plate (B). Expression continues in the endoderm at stage 26 (C). A section through a stage 26 embryo shows that *edd* is also expressed in the notochord (D). d = dorsal; v = ventral; a = anterior; p = posterior; no = notochord; np = neural plate; nt = neural tube; so = somite; endo = endoderm.

Figure 5.3 Comparison of the expression pattern of *edd* and *chd* in gastrulae

Chd transcripts are localised exclusively to the DMZ at stage 10/10.5 (A), while *edd* transcripts appear enriched in this region but also extend throughout the marginal zone. Note the speckled pattern of *edd* expression (B). Bisecting the embryo through the DMZ shows cells expressing *chd* in the mesoderm at the dorsal lip and in cells migrating towards the animal pole (C). These migrating cells will later form the notochord. *Edd* expression is seen in deep lying endoderm on the dorsal side and expressed in some individual endodermal cells ventrally (D). Note that *edd* transcripts appear localised to the nucleus. *Edd* expression overlaps with that of *chd* in the involuting mesoderm, consistent with its later expression in the notochord. Panels E and F show two halves of a bisected embryo stained for either *chd* (E) or *edd* (F). d = dorsal; v = ventral.

Figure 5.2

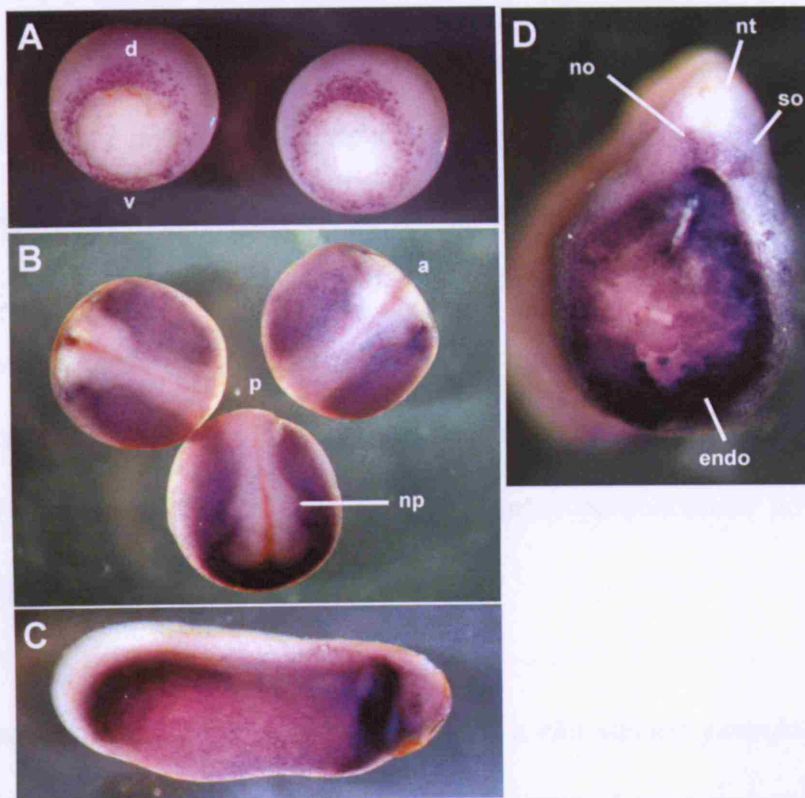
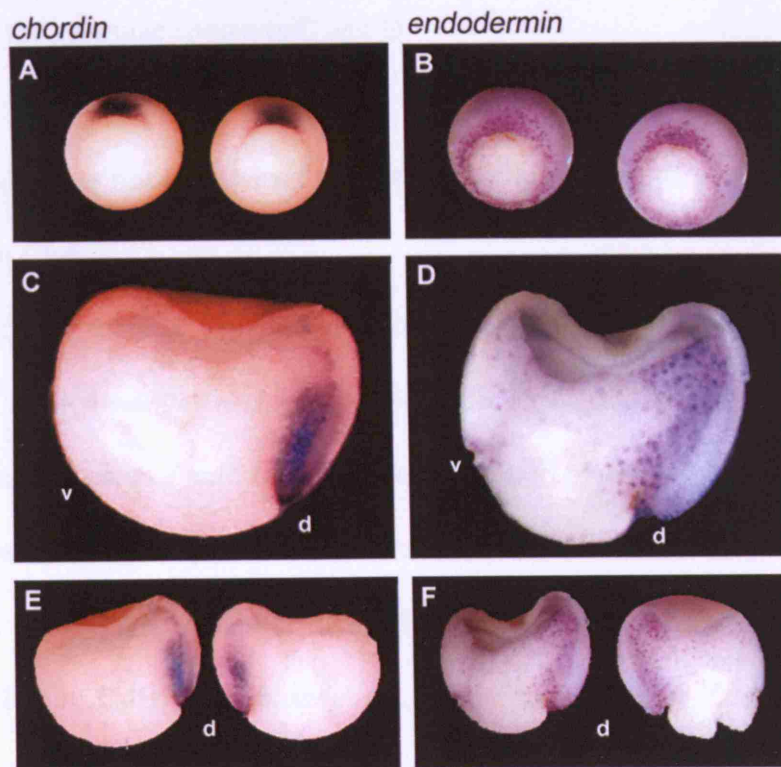


Figure 5.3



A cross-section through the DMZ shows *edd* transcripts in deep lying cells which will go on to form endoderm, again in a diffuse speckled manner (Figure 5.3D). Expression of *edd* appears to overlap with that of *chd* in the involuting mesoderm that will contribute to the notochord, this is consistent with later expression of *edd* in this structure. The yolky endoderm cells are quite large in comparison to the outer ectoderm or mesodermal cells of the marginal zone and the intense staining of *edd* expression in the middle of these endoderm cells indicates to us that transcripts are localised around, or in, the nucleus. This cannot be confirmed however without closer, higher resolution microscopy.

In comparison to *edd*, a cross-section through a *chd* stained gastrula (Figure 5.3C) revealed no staining in the endoderm and enrichment of transcripts in the mesoderm and tissue involuting though the dorsal blastopore lip, cells that will form the notochord and paraxial mesoderm. Figure 5.3E and F show the two corresponding halves of the gastrulae shown in C and D.

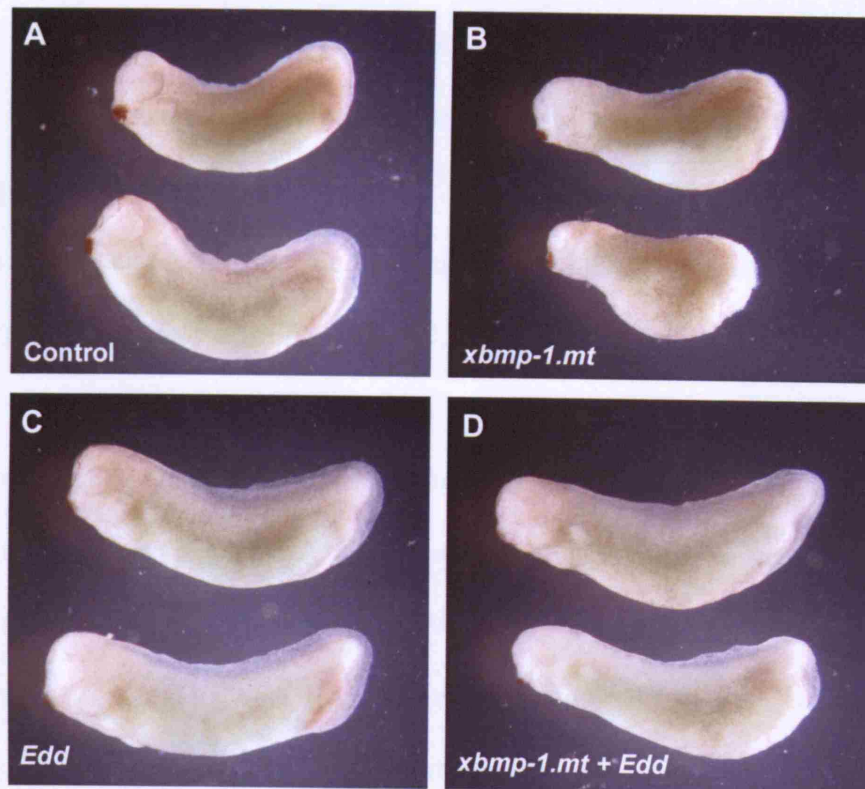
5.2.2 Edd does not counteract the effect of *xbmp-1* over-expression, or its action on Chd.

To investigate whether Edd has the ability to inhibit members of the Tld family, synthetic capped *edd* mRNA was injected into 2 cell stage embryos in the presence and absence of *xbmp-1* mRNA. Ectopic expression of *xbmp-1* throughout the whole embryo results in the expansion of ventral cell fates due to increased cleavage of Chd, an effect that causes reduction of head structures and an enlargement of posterior regions (Figure 5.4B) (Goodman et al., 1998). Inhibition of XBMP-1 on the other hand results in the opposite

Figure 5.4 Edd does not inhibit the ventralising activity of BMP-1

Embryos were injected at the two cell stage with either 1.6ng of mRNA for *xbmp1.mt* (B), 1.6ng mRNA for *edd* (C), or 1.6ng *edd* and 1.6ng *xbmp-1.mt* (D). Injected embryos were cultured until uninjected sibling embryos (A) had reached stage 29/30. Embryos injected with *xbmp1.mt* were ventralised in 37/41 cases this was not counteracted by co-injection of *edd* mRNA which induced ventralisation in 29/29 cases. No ventralisation was seen in either control embryos or *edd* injected embryos (0/52 and 0/32 respectively). The embryos shown are representative of the phenotype obtained in two independent sets of injections.

Figure 5.4



effect dorsalising the embryo, with larger heads and reduced posterior structures (Wardle et al., 1999). If Edd is a strong inhibitor of XBMP-1 then ectopic expression of *edd* should dorsalise *Xenopus* embryos. However, over-expression of *edd* mRNA results in no detectable phenotypic defect when compared to uninjected siblings (Figure 5.4A & C). Similarly, co-injection of *edd* with *xbmp-1* mRNA should prevent *xbmp-1* induced ventralisation of the embryo. Injection of 1.6ng *xbmp-1* mRNA and 1.6ng *edd* mRNA resulted in a phenotype similar to that of *xbmp-1* injection alone but with perhaps less severity (Figure 5.4D). This indicated that Edd might have a small effect counteracting XBMP-1 and although this could be explained by discrepancies in the amount of *xbmp-1* mRNA injected, it was decided to investigate this further.

The ventralising ability of XBMP-1, when over-expressed in the VMZ, is relatively weak (Goodman et al., 1998). It will however counteract the formation of secondary axes induced by ectopic expression of *chd* in the same region (Wardle et al., 1999). We therefore tested whether Edd can counteract the anti-Chd activity of XBMP-1. Embryos were injected into two ventral-vegetal cells at the 4 to 8 cell stage with either *chd*, *chd* and *xbmp-1* or *chd* and *edd*, and triple injected with *chd*, *xbmp-1* and *edd* mRNA. 0.8ng of mRNA was injected for each gene. Ventral injection of *chd* alone resulted in secondary axes in 63% of embryos (25/40) (Figure 5.5B), while co-injection of *chd* with *xbmp-1* resulted in secondary axes in 0% of embryos (0/43) (Figure 5.5C) and in some cases the embryos were weakly ventralised. Co-expression of *chd* with *edd* resulted in secondary axes in 56% of cases (23/41) (Figure. 5.5D) showing that Edd does not have an effect on Chd. Triple injection of *chd*, *xbmp-1* and *edd* resulted in one secondary axis

Figure 5.5 Edd does not prevent BMP-1 from blocking the secondary axis inducing activity of Chd

Embryos were injected into two ventro-vegetal cells at the 4 to 8 cell stage with either 0.8ng *chd* mRNA (B), 0.8ng *xbmp-1* and 0.8ng *chd* mRNA (C), 0.8ng *edd* and 0.8ng *chd* mRNA (D), or 0.8ng *chd*, 0.8ng *xbmp-1* and 0.8ng *edd* mRNA (E). Injected embryos were cultured until uninjected siblings (A) reached stage 25/26 and assessed for secondary axis formation (arrowheads). A sample of 4 embryos from each set of injections is shown in each panel. Note that BMP-1, but not Edd, blocks the ability of Chd to induce a secondary axis, and that Edd does not inhibit BMP-1 in this assay. Also, note that some of the embryos in panels C and E have a BMP-1 phenotype.

Figure 5.6 Edd does not prevent BMP-1 from blocking the dorsalising activity of Chd in VMZs

Embryos were injected ventro-vegetally at the 4 to 8 cell stage with either 0.8ng *chd* mRNA (B), 0.8ng *chd* + 0.8ng *xbmp-1* mRNA (C), 0.8ng *chd* + 0.8ng *edd* mRNA (D), or 0.8ng *chd* + 0.8ng *xbmp-1* + 0.8ng *edd* mRNA (E). VMZ's were isolated at stage 10/10.5 from each batch of embryos, including uninjected controls (A), and incubated until stage 20. to assess for elongation. Note that Chd induces a pronounced elongation of the VMZ and that this is inhibited by BMP-1 but not Edd. Also, note that Edd has no effect on BMP-1 in this assay.

Figure 5.5

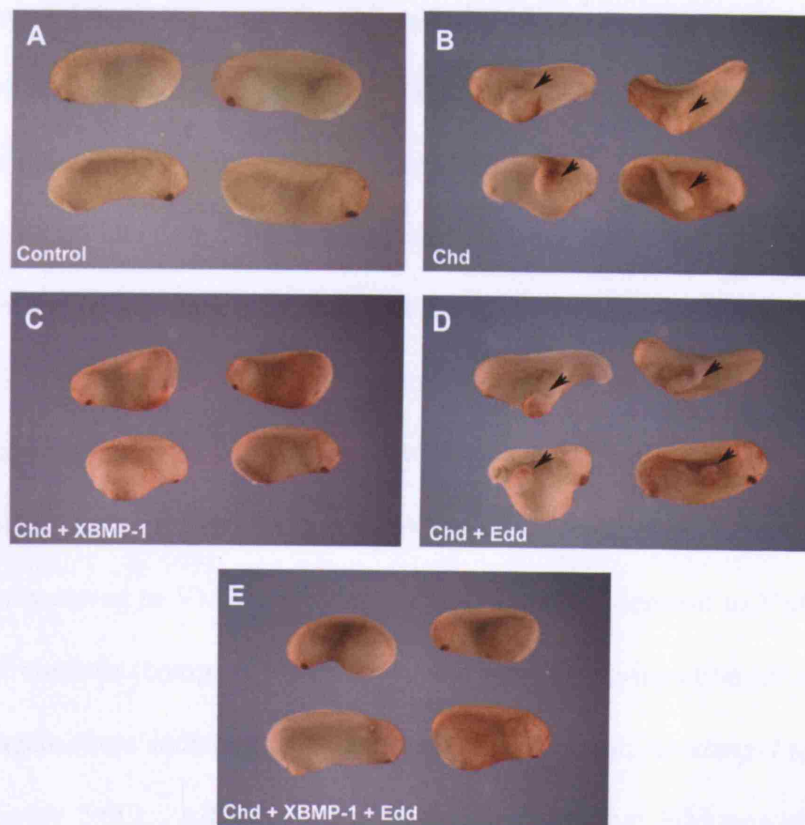
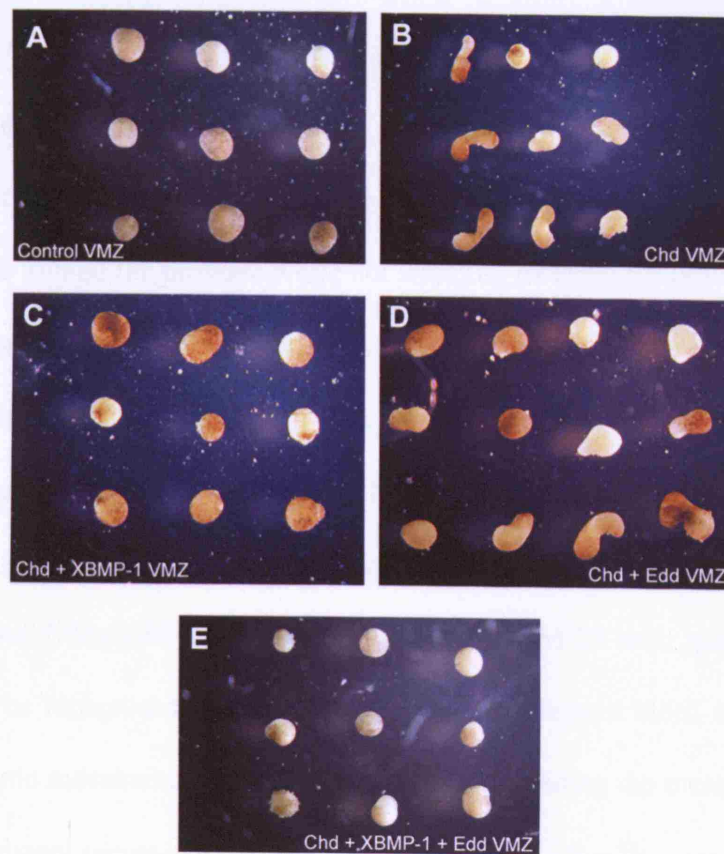


Figure 5.6



being formed 2% (1/49), however, the majority of embryos were identical to those injected with *chd* and *xbmp-1* (compare Figure 5.5C & E). No secondary axes were formed in uninjected sibling controls ((0/43) Figure 5.5A). In parallel to looking at secondary axis induction, VMZs were also cut from a number of ventro-vegetally injected embryos at stage 10/10.5 and cultured in higher strength NAM salts (NAM/2). Ventral expression of *chd* resulted in elongation of excised VMZs as a result of convergent extension movements (Figure 5.6B), similar to those found when DMZ explants are cultured in isolation. When *xbmp-1* was co-expressed with *chd* no elongation occurred in VMZ fragments, which appeared identical to VMZs cut from uninjected controls (compare Figure 5.6A and C). Triple injection of *xbmp-1*, *chd*, and *edd* again were indistinguishable from either control, or *xbmp-1+chd* injected VMZs (Figure 5.6E). All three experiments indicated that Edd was not inhibiting XBMP-1 activity in the embryo.

5.2.3 Edd does not bind XBMP-1

Edd does not counteract XBMP-1 activity in the whole embryo but that is not to say Edd is unable to bind the protease while not affecting its catalytic activity. In the case of α -2M, trapped proteases still retain their activity allowing some substrates to be cleaved (Sottrup-Jensen et al., 1989). We therefore wondered if this was the case for Edd. To investigate whether Edd physically binds XBMP-1, embryos were injected at the 2 cell stage with either 1.6ng of myc-tagged *xbmp-1* (*xbmp-1.mt*) mRNA or 1.6ng *xbmp-1.mt* and 1.6ng *edd* mRNA and cultured in NAM/20 until gastrulation. The myc tag can be recognised by the 9E10 antibody on Western blots, therefore if Edd acts in the same mechanism as α -2M, i.e. covalently binding the protease, one would expect an apparent increase in the size of XBMP-1.MT when the assay is done under

non-reducing conditions. XBMP-1.MT has a molecular weight of 105 kDa (Figure 5.7), but includes a pro-peptide that is removed to yield an activated protease of approximately 95 kDa. It is this activated protein that would be bound by Edd. In the presence of Edd there is no apparent size increase of XBMP-1.MT and the 95 kDa band is present in both lanes (Figure 5.7 compare lanes B and C). This indicates that Edd does not bind XBMP-1.

5.2.4 α -2 macroglobulin does not inhibit cleavage of Chd in Xlr conditioned media

While Edd does not appear to bind nor inhibit XBMP-1 we investigated whether α -2M itself inhibits another Tld-related protease, Xlr. Xlr is closely related to XBMP-1, elicits the same effects on Chd and ventralises *Xenopus* embryos when over-expressed (Dale et al., 2002). Xlr was chosen for this particular experiment due to the availability of active conditioned media for the protease that could be combined with commercially available α -2M protein (Sigma). Myc tagged Chd protein was incubated for 14 hrs at 30° C in Xlr conditioned media (Dale et al., 2002) in both the presence and absence of excess α -2M. In this experiment α -2M did not affect the ability of Xlr to cleave Chd (Figure 5.8 lane B and C). Chd and Xlr were also incubated in the presence of 1,10 Phenanthroline (1,10PT), a potent chemical inhibitor of zinc metalloproteases that blocked cleavage of Chd by Xlr (Figure 5.8 lane D). This experiment demonstrates that α -2M is not an inhibitor of Xlr.

Figure 5.7 Edd does not bind to BMP-1

Embryos injected with either 1.6ng of *xbmp-1.mt* mRNA or 1.6ng of *xbmp-1.mt* + 1.6ng *edd* mRNA were collected at stage 35 and homogenised. One embryo's worth was subjected to Western blot analysis under non-reducing conditions, using the 9E10 antibody to detect the myc tagged XBMP-1. The majority of XBMP-1.MT detected was 105 kDa in size, with a small fraction detected at 95 kDa (Lane B). The lower band is likely to be activated XBMP-1, after pro-peptide removal. There was no difference in the size of XBMP-1.MT from embryos co-injected with Edd (Lane C), indicating that Edd does not bind to XBMP-1. Uninjected control embryos were used as a negative control (Lane A).

Figure 5.8 α -2M does not inhibit Xlr

Myc tagged Chd protein was incubated with either Control media, or media conditioned for Xlr in the presence and absence of human α -2M. Chd was cleaved by Xlr conditioned media (Lane B), but not by control media (Lane A). Cleavage of Chd by Xlr was inhibited by the specific inhibitor of Zn^{2+} metalloproteases 1,10 Phenanthroline (Lane D), but not by α -2M (Lane C). This result is representative of two independent experiments.

Figure 5.7

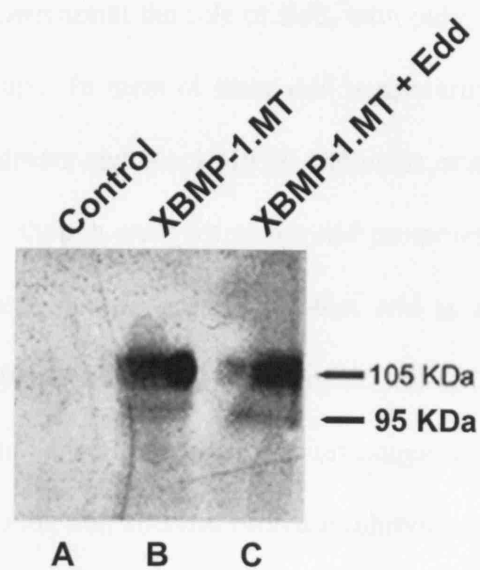
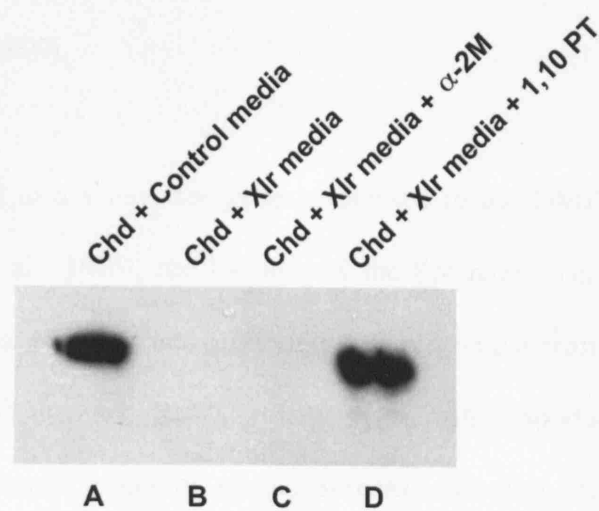


Figure 5.8



5.3 Discussion

5.3.1 Edd in *Xenopus* development

Currently very little is known about the role of Edd, with only 10 papers having been published in the last 9 years. In most of these *edd* is primarily used as a marker of endoderm formation (Chalmers and Slack, 1998; Clements et al., 2003; Zhang et al., 2004; Zhang et al., 1998), though analyses of the *edd* promoter has shown it to have binding sites for Sox transcription factors and that *edd* is a direct target of the endodermal transcription factor XSox17 (Ahmed et al., 2004; Clements et al., 2003). In this chapter I have taken a different approach to investigating Edd function, starting with its close homology to the non-specific protease inhibitor α -2M. Our hypothesis was that Edd would also be a protease inhibitor that is required to regulate proteolytic activity in the endoderm and notochord of early *Xenopus* embryos. More specifically, we have explored the possibility that Edd is an inhibitor of members of the Tld family of metalloproteases, which are believed to be important regulators of BMP signalling in the mesoderm and endoderm of early *Xenopus* embryos (Dale and Jones, 1999; Dale and Wardle, 1999).

Edd was identified in a screen for genes expressed in the DMZ of early *Xenopus* gastrulae (Sasai et al., 1996), the location of the Spemann organiser an important signalling centre that plays a crucial role in dorsal-ventral patterning of all three germ layers (Harland and Gerhart, 1997). Many genes with important roles in dorsal-ventral patterning were found in similar screens, including the BMP inhibitors, noggin, chordin and cerberus (Bouwmeester et al., 1996; Sasai et al., 1994; Smith and Harland, 1992). Most of these genes were found exclusively expressed in the mesoderm and/or endoderm of the organiser and misexpression disturbed dorsal-

ventral patterning. In contrast, *edd* was found more widely expressed in the marginal zone of early gastrulae, although enriched on the dorsal side (Sasai et al., 1996). My initial comparison of *edd* expression with that of Sasai *et al* confirmed the latter's result, with expression in the marginal zone which was enriched in the DMZ and later localises to the endoderm and notochord. A closer inspection of expression at gastrulation showed that cells expressing *edd* appeared in a speckled pattern and transcripts seemed to be located to the centre of cells, this was most apparent in the large yolky vegetal pole/endodermal cells. This might indicate that they are located in or around the nucleus but higher magnification microscopy or double *in situ* with a probe for nuclear localised transcripts would be required to confirm this. We wished to determine if Edd could inhibit members of the Tld family and protect Chd in the DMZ, therefore a closer comparison of the expression of each gene was made at gastrulation. Bisected embryos showed that *edd* transcripts were in fact on the superficial layer of epidermis above the DMZ, while *chd* transcripts were located deeper in the mesoderm. Sasai *et al.* reported that the *edd* expressing superficial cells were fated to become endoderm, consistent with its pan-endodermal expression, we however are unable to confirm this. *Edd* expression does overlap with that of *chd* in the involuting mesoderm that is likely to become notochordal and paraxial mesoderm and is therefore consistent with its later presence in the notochord proper. *Chd* expression is specifically localised to the DMZ and area of involuting mesoderm and, while partially overlapping with this expression, bisected embryos showed that *edd* was more widely expressed in the endoderm than previously described at this stage. *Edd* expressing cells can be seen throughout the endoderm again located to the centre of cells. While this could be the result of, what Wardle and Smith (2004) term as "rogue" expressing cells, their prevalence leads me to believe that this is unlikely. A

more plausible explanation is that *edd* expression is marking the initial specification of these cells to endoderm and a slightly later stage might show more cells expressing *edd* transcripts in this manner.

The results presented in this chapter indicate that Edd is not an inhibitor of the Tlds. Co-expression of mRNA for *edd* with *xbmp1* was unable to prevent the ventralisation of mesoderm resulting in a reduced anterior and enlarged posterior of tail bud stage embryos, nor was it able to counteract the chordinase activity of XBMP-1 in either VMZ or secondary axis induction assays. Additionally I also tried to determine if Edd could bind XBMP-1 but not necessarily inhibit the protease, as proteases bound by α -2M often remain active (Sottrup-Jensen et al., 1989). This was not the case, although to be sure a co-precipitation experiment would be required but we did not believe this necessary, as it was now apparent Edd was unlikely to be an inhibitor of XBMP-1. Finally, using conditioned media for Xlr, closely related to XBMP-1, I have investigated the possibility of α -2M itself acting as an inhibitor of the Tlds. α -2M failed to inhibit cleavage of Chd by Xlr and I am therefore confident that neither Edd nor α -2M are inhibitors of the Tlds.

5.3.2 Potential roles of Edd and α -2M as protease inhibitors during embryogenesis

The results presented in this chapter show that Edd does not inhibit, nor physically bind, XBMP-1, nor does it result in any other discernible embryonic defects. In contrast to the few papers published on Edd, α -2M has been extensively studied for its action as a protease inhibitor but its role in developing embryos is equally unclear. In mouse embryos distribution of mRNA transcripts for α -2M is localised to the liver

from Day 13 but undetected before this time (Lorent et al., 1994) and null mutants for α -2M do not have any developmental abnormalities with all offspring being viable (Umans et al., 1995; Umans et al., 1999). This suggests the protein has no role in embryonic patterning, or that other proteins compensate for α -2M functions. However, the localisation of α -2M exclusively to the embryonic liver indicates that it is not a true functional homolog of Edd, which is more ubiquitously expressed and appears earlier in development and before any liver precursors are present.

Despite this, the close homology between the two proteins cannot be ignored. The presence of a thio-ester group in both proteins, important for covalently binding molecules within α -2M inhibitory complexes, indicates that Edd may well bind proteins in a similar fashion to α -2M. A literature search for proteases present during *Xenopus* gastrulation identifies a number of enzymes including the serine proteases Furin/XPACE4 (Birsoy et al., 2005), Xesp-1, Xesp-2 and XMP-SP1 (Yamada et al., 2000). There are also metalloproteases from the ADAM family such as ADAM-13 (Alfandari et al., 1997) and MMP family including XMMP, MMP-14 and MMP-15 (Harrison et al., 2004; Yang et al., 1997). α -2M has been shown to inhibit members of all these classes of proteases (Sottrup-Jensen et al., 1989) and it is possible that Edd may inhibit these, or indeed yet to be identified enzymes. Of those mentioned above however, XPACE4, although found in endodermal cells early in gastrulation, is located to intracellular compartments (Birsoy et al., 2005) and may not come into contact with Edd, and the serine proteases Xesp-1, -2 and XMP-SP1, for which no specific function has been allied as yet, appear to be ubiquitously expressed (Yamada et al., 2000). Although in the case of the later three enzymes Edd could function to provide to local inhibition in the marginal zone if it acts as a protease inhibitor.

5.3.3 Other potential roles for Edd

It is possible however that Edd does not act as a protease inhibitor at all. In addition to proteases, α -2M been reported to bind a number of growth factors including TGF β -1, TGF β -2, FGF-2 and Activin (Danielpour and Sporn, 1990; Dennis et al., 1989; Vaughan and Vale, 1993). The mechanism by which this occurs is still unclear, but is likely to occur via hydrophobic sequences and different acidic residues in the mature protein (Arandjelovic et al., 2003). In all cases the signalling molecule is unable to function via its normal receptor pathway. There are a large number of different growth factors present during *Xenopus* gastrulation, including members of both the TGF- β and FGF families (Hill, 2001; Isaacs, 1997). These molecules have multiple roles in *Xenopus* development, including mesodermal and neural development, and changes in their activity cause major disruptions to embryonic patterning and development. Interestingly, at gastrulation *eFGF* is expressed in a very similar pattern to that of *edd* (Isaacs et al., 1995), however, modulation of eFGF activity severely affects mesoderm formation. Over-expressing *edd* on the other hand has no discernible effect on *Xenopus* development, one would expect that even a small change to signalling levels of any growth factor important during gastrulation would be revealed by some patterning defect. Thus indicating that Edd has little effect on signalling by these growth factors during this time.

5.3.4 Summary

Edd is a *Xenopus* protein closely related to α -2M a non-specific inhibitor of extracellular proteases. In this chapter I investigated whether Edd could act in the capacity of an inhibitor for XBMP-1, unfortunately this was not the case. Although in this chapter I have attempted to identify an inhibitor of Tld proteases to attenuate their

activity in the DMZ a number of alternative explanations can be proposed as to why an inhibitor may not be necessary. Firstly, the presence of XBMP-1 and Xld in the DMZ, as indeed the entire marginal zone, was detected using techniques identifying mRNA transcripts (Goodman et al., 1998) and therefore we do not know the protein distribution of these two proteases. It is entirely possible therefore that within the DMZ these mRNAs are not actually translated. Secondly, we have seen in Chapter 4 that certain CUB domains play a role in the secretion and of Tlds. This raises the possibility that an internal mechanism affecting the CUB domains in the secretory pathway may preclude secretion and/or their activation in the DMZ. Finally, little is known about the protein folding of the Tlds and there may exist an unknown mechanism that prevents correct folding of the proteases thereby attenuating their function.

Subsequent to finishing this thesis an endogenous inhibitor of the Tld protease Xlr has been identified. Lee et al., (2006) have discovered that Sizzled, a ventrally expressed secreted Frizzled like molecule, competitively inhibits Xlr preventing the proteases cleavage of Chd. A similar molecule may therefore exist in the DMZ to attenuate XBMP-1 and Xld activity.

Chapter 6

Identification of potential substrates for the *Xenopus* Tlds

6.1 Introduction

6.1.1 Substrates of the Tlds

With the exception of Chd, few of the known mammalian substrates for the Tlds have been described in *Xenopus*. Although not formally determined as substrates for the *Xenopus* Tlds, only *proCol1 α 1* and *proCol2 α 1* have been described during early *Xenopus* development (Goto et al., 2000; Su et al., 1991). In mammals, both have their C-terminal pro-peptide removed by BMP-1/PCP, which allows the formation of insoluble collagen fibres into mature fibrils, mediated by the amine oxidase, Lox (Csiszar, 2001; Kagan, 1986; Prockop and Kivirikko, 1995) (also see Chapter 3). With the exception of the most recently described *proCol24 α 1* and *proCol27 α 1* (Koch et al., 2003; Pace et al., 2003), most fibrillar procollagens have been shown to be substrates of BMP-1. This Tld family member removes the C-terminal pro-peptide from *proCol3 α 1*, both the N- and C-terminal pro-peptides of *proCol5 α 1* and *proCol5 α 2*, and the N-terminal pro-peptide of *proCol11 α 1* (Imamura et al., 1998; Kessler et al., 2001; Medeck et al., 2003; Unsold et al., 2002). A quick analysis of the protein sequences for both *proCol24 α 1* and *proCol27 α 1* identifies a number of potential sites for Tld cleavage. In *proCol27 α 1* there is a conserved aspartic acid residue near the beginning of the C-terminal pro-peptide (aspartic acid is always found in the P1 position of BMP-1 cleavage sites), indicating that this collagen might be a substrate for BMP-1. Although this residue has not been conserved in *proCol24 α 1*, a number of aspartic acids are found close to the beginning of the C-terminal peptide. Without sequencing these proteins after incubation with BMP-

1/PCP incubation, the role of the enzyme can only be speculation. Of the non-fibrillar collagens only procollagen type VII, involved in connecting skin to the basement membranes, is processed by the Tlds, though mTII-2 is not as efficient as the others (Rattenholl et al., 2002).

Two other factors, the SLRPs biglycan (Bgn) and osteoglycin (Ogn), appear to be involved in collagen formation and are also processed by members of the Tld family (Ge et al., 2004; Scott et al., 2000). Bgn associates with side chains of $\text{Coll}\alpha 1$, regulating fibril formation, while Ogn, whose mode of action is unclear at present, also regulates collagen fibril assembly, a role enhanced by proteolytic cleavage with BMP-1. Both Bgn and Ogn are processed by mammalian BMP-1, but only Ogn is processed by mTld and mTII-1 as well. mTII-2 is not effective against either substrate. Interestingly, Lox, which mediates collagen fibre formation, itself is secreted in an inactive latent state that is activated by Tlds, which remove the N-terminal pro-peptide, yielding a catalytically active enzyme (Panchenko et al., 1996; Uzel et al., 2001). A number of enzymes related to Lox have been identified and of these Loxl-1 is also reportedly activated by BMP-1 in mammalian systems (Borel et al., 2001).

The action of the Tlds in processing ECM proteins is not confined to collagen formation, however. BMP-1 processes the γ -2 chain of laminin-5 and has been implicated in the same role for the α -3 chain (Amano et al., 2000). In doing so it is thought BMP-1 controls the interaction of C-terminal globular domains (LG) in these laminin macromolecules with proteins in the basal lamina. The Tlds also remove a C-terminal domain of the heparin sulphate proteoglycan perlecan (Gonzalez et al.,

2005). The domain, known as endorepellin, contains a series of LG-like repeats and is thought to elicit its own angiostatic effect. All the mammalian Tlds can remove this module from perlecan, but again mTll-2 is the least effective. Finally, BMP-1, mTld, mTll-1 and mTll-2 all processes the murine negative muscle growth regulator, and TGF- β family member, myostatin from a latent to an active form, although this time mTld is not as efficient as the others, demonstrating a further more direct role in growth factor control (Wolfman et al., 2003).

In *Xenopus* little sequence data is available for the fibrillar collagens, apart from the already described proCol1 α 1 and proCol2 α 1. Both of these *Xenopus* proteins contain conserved Tld cleavage sites, but have yet to be formally demonstrated as substrates for these proteases. In Chapter 3 we also saw that *Xenopus* Lox contains a conserved BMP-1 cleavage site, but this is not the case for XLoxl-1. Complete sequences for both *Xenopus* Bgn (XBgn) and Ogn (Ogn) have been deposited in the sequence databases (Acc: AB037269 and BC043788), but only XBgn appears to contain a Tld cleavage site. The nucleotide databases contain no *Xenopus* sequences for either *myostatin* or *laminin-5 γ -2*, but does contain an EST annotated as *laminin-5 α -3*.

6.1.2 Aims

To further understand the role of the Tlds in *Xenopus* embryogenesis we wished to identify putative substrates of the proteases expressed early in development. Using the EST databases of NCBI and NIBB a search was made for *Xenopus* ESTs that displayed significant homology to, or were annotated as known substrates of, the mammalian Tlds. This search identified nine different genes, including three members of the Lox family already described in Chapter 3. Of the remaining six

ESTs, four were identified as fibrillar procollagens *proCol3 α 1*, *proCol5 α 1*, *proCol5 α 2*, and *proColl1 α 1*, a fifth as Bgn, and the sixth was annotated as being the α -3 chain of laminin-5. This chapter describes the expression of each of these genes during early *Xenopus* development.

6.2 Results

6.2.1 Fibrillar procollagens expressed in early *Xenopus* development

The fibrillar collagens proCol3 α 1, proCol5 α 1, proCol5 α 2 and proCol11 α 1 are all known substrates of BMP-1/PCP (Imamura et al., 1998; Kessler et al., 2001; Medeck et al., 2003; Unsold et al., 2002). A search of the EST databases revealed a number of *Xenopus laevis* clones closely related to each of the afore-mentioned genes (for accession numbers please see Table 2.1 in Chapter 2). Each clone was sequenced from each end to confirm the identity of each EST, to determine its orientation for transcription of *in situ* hybridisation probes, and for designing RT-PCR primers. A nucleotide alignment for each sequence was made to their human homologues and these are presented as appendices (Appendix D1, D2, D3 & D4).

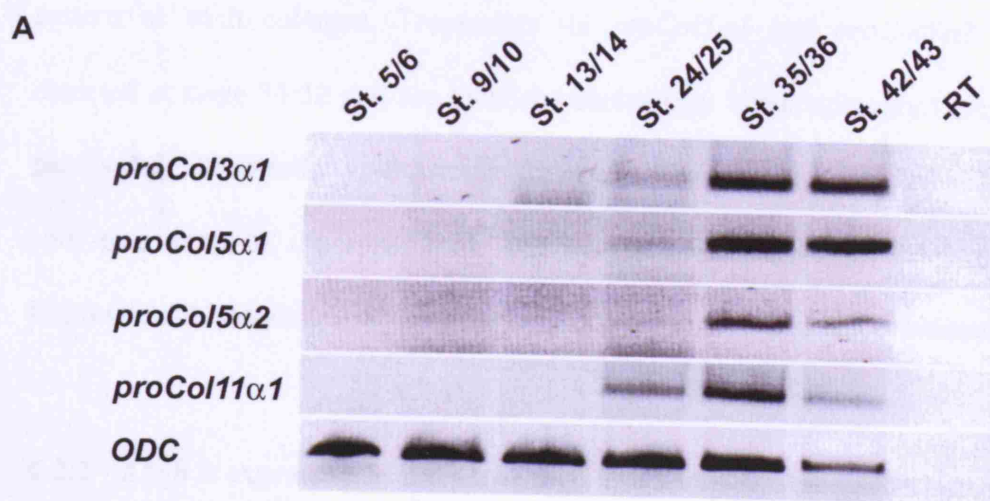
The temporal expression pattern of each clone was determined by RT-PCR, using total RNA extracted from six different stages of development (Figure 6.1A). *proCol3 α 1*, *proCol5 α 2*, and *proCol11 α 1* were first detected at stage 24/25, while *proCol5 α 1* was faintly detected at stage 13/14. All genes increase levels of expression between stage 24/25 and stage 35/36. A drop in the levels of *ODC* at stage 42/43 makes it difficult to fully determine how expression progresses into this stage. However, there is a corresponding drop of expression for *proCol5 α 2* and *proCol11 α 1*, suggesting transcript levels of these two genes may remain the same, while levels of *proCol3 α 1* and *proCol5 α 1* do not display this drop leading me to believe expression of these two genes will rise between stage 35/36 and 42/43, similar to the rise observed for *proCol1 α 1* and *proCol2 α 1* (see Figure 3.3).

Figure 6.1 Expression of fibrillar procollagens in early *Xenopus* development

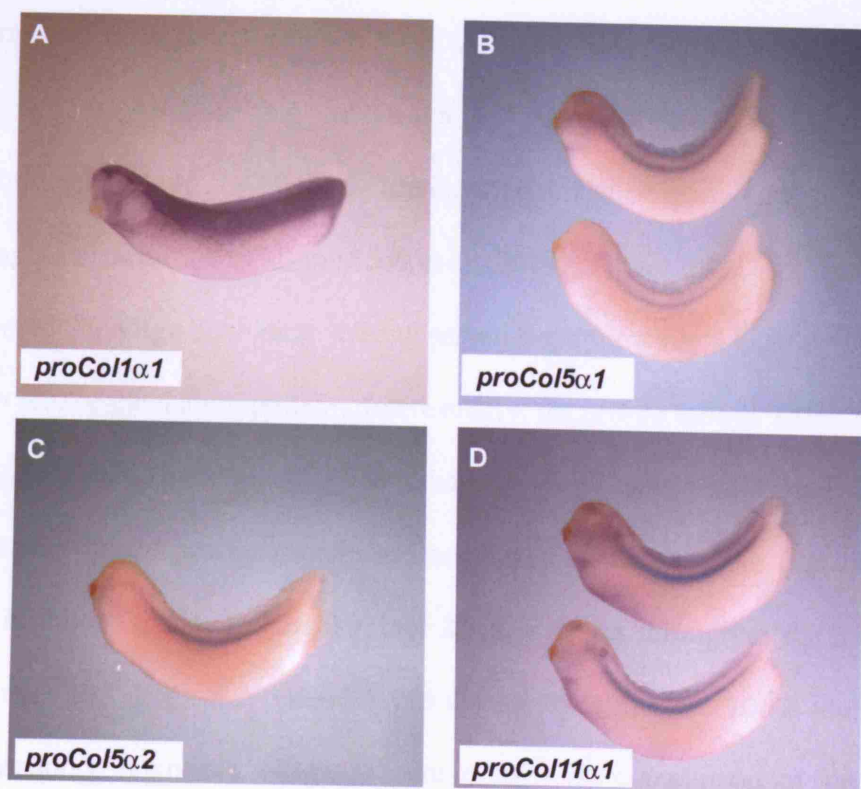
A. Total RNA was extracted from six different stages of *Xenopus* development and subjected to RT-PCR using specific primers for *proCol3 α 1*, *proCol5 α 1* and α 2, and *proCol11 α 1*. Primers for *ornithine decarboxylase* (*ODC*) were used to control for differences in cDNA levels between different samples. Total RNA from stage 24/25 embryos, minus reverse transcriptase (-RT), was used as a negative control. Transcripts for *proCol5 α 1* were the first to be detected, with a faint signal observed in the stage 13/14 lane. For the other procollagens, transcripts were first detected at stage 24/25, although the signal for *proCol5 α 2* is very faint. Transcript levels for all four genes had increased significantly by stage 35/36. Unfortunately, levels of *ODC* dropped at stage 42/43, making it difficult to accurately determine what happens to the *collagens*. However, from the presented data I believe that transcript levels for *proCol3 α 1* and *proCol5 α 1* probably increase between these stages, while transcripts for *proCol5 α 2* probably remain level. Transcript levels for *proCol11 α 1* either remain level or drop slightly. RNA from stage 24/25 embryos, minus reverse transcriptase, was used as a negative control. The result shown is representative of at least three independent experiments.

B. Whole mount *in situ* hybridisation of stage 28-32 embryos with antisense DIG labelled probes for *proColl α 1* (A), *proCol5 α 1* (B), *proCol5 α 2* (C), and *proColl11 α 1* (D). Transcripts for *proColl α 1* are found in the dorsal epidermis, while transcripts for *proCol5 α 1*, *proCol5 α 2*, and *proColl11 α 1* are localised to the notochord. I have been unable to obtain an *in situ* pattern for *proCol3 α 1*.

Figure 6.1



B



Whole mount *in situ* hybridisation was used to determine the spatial expression pattern of each collagen. Transcripts for *proCol5α1* and *proCol5α2* were first detected at stage 31/32 and are localised exclusively to the notochord (Figure 6.1B, panels B & C). Similarly, transcripts for *proColl1α1* are first detected at stage 31/32 and are found in the notochord, otic vesicle, and fore-brain (Figure 6.1B d). Expression at later stages of development remains identical to that presented here.

6.2.2 *Xbgn* is expressed in the notochord and somites after gastrulation

Searching the NCBI EST database we identified an EST (BG729865) displaying close homology to *Xbgn*, here after referred to as *Xbgn-EST*. The clone contained a 1941 bp fragment that at the nucleotide level displayed 99% homology to that of the previously published sequence for *Xbgn* (Acc: AB037269), in the overlapping region. *Xbgn* is 2960 bp long and encodes a 368 amino acid protein (coded between nucleotides 266 and 1372) 77% identical to hBgn (Figure 6.2A). The BMP-1 cleavage site identified in hBgn (Scott et al., 2000) (Figure 6.2A highlighted in red) is conserved in XBgn including the conserved aspartic acid, which will be the first amino acid of the mature protein. Interestingly, the pro-peptide of XBgn contains two cysteine residues (highlighted in blue) that are not present in the human sequence, all 7 cysteines in the mature protein are conserved however. The first nucleotide of *Xbgn-EST* corresponds to base 1021 of *Xbgn* and runs through to the 3' poly A tail (Appendix E). *Xbgn-EST* encodes 116 amino acids that are 100% identical to the corresponding region of XBgn (Figure 6.2B). A comparison of the nucleotide sequence within the 3' UTR of *Xbgn* and *Xbgn-EST* reveals that the two sequences are identical with the exception of a 6 nucleotide insertion in

Figure 6.2 Amino acid sequence of XBgn and Xbgn-EST

A. Amino acid alignment of human and *Xenopus* biglycan (hBgn and XBgn). Both proteins are 368 amino acids in length and are 77% identical to each other. The cleavage site for BMP-1 is indicated (boxed and red letters), with cleavage in the human sequence occurring between amino acids 36 and 37. Seven conserved cysteine in the mature protein are highlighted (in blue) as are two cysteines found in the signal peptide of XBgn (in purple).

B. Alignment of the deduced amino acid sequence encoded by *Xbgn-EST* with the overlapping sequence of XBgn. In the region of overlap, XBgn-EST and XBgn are 100% identical.

Figure 6.2

A

```

hBgn          1 MWPLWRLVSL LALSQALPFEQRGFWDFTLDDGPFMMNDEEASGADTSGVL 50
Xbgn AB037269 1 MKVLLLLCSCILVIHALPFEQRGFWDFSMDDGMAMMKDEEASGVGPIPTE 50
               * * * * . . .*****.*** ** *****

hBgn          51 DPDSVTPTYSAMCPFGCHCHLRVVQCS DLGLKSVPKEISPDITLLDLQNN 100
Xbgn AB037269 51 SIPDVGLPPMDLCPFGCQCHLRVVQCS DLGLTSIPKNLPKDTLLDLQNN 100
               * .*****.***** *.*. *****

hBgn          101 DISELRKDDFKGLQHL YALVLVNNKISKIHEKAFSPLRKLQKLYISK NHL 150
Xbgn AB037269 101 KITEIKKDDFKGLTNLYALVI VNNKISKINEKA FEPLQKMQLYISKNNL 150
               *.***.*****.*****.***** **.*.*****.*

hBgn          151 VEIPPNLPSLVELRIHDNRIRKVPKGVFSGLRNMNCIEMGGNPLENSGF 200
Xbgn AB037269 151 EEIPKNLPKSLVELRIHENKIKKVPKGVFSGLKNMNCIEMGGNPLENGGI 200
               *** ** *****.*.*.*****.***** **

hBgn          201 EPGAFDGLKLN YLRIS EAKLTGIPKDLPETLNLHLDHNKIQAIELEDLL 250
Xbgn AB037269 201 EAGAFDGLKLN YLRVSEAKLSGIPKGLPSTLNLHLDHNKIQAIEKEDLS 250
               * *****.*****.*** ** *****.***** **

hBgn          251 RYSKLYRLGLGHNQIRMIENGSLSF LPTLRELHLDNNKLARVPSGLPDLK 300
Xbgn AB037269 251 QYASLYRLGLGHNNIRMIENGSLSF MPVLRRELHLDNNKLSKVPPGLPDMK 300
               .*.*****.*****.* *****.*** **.*.

hBgn          301 LLQVVYLHSNNITQVG VNDFCPMGFVKRAYNGISLFNNPVPYWEVQPA 350
Xbgn AB037269 301 LLQVVYLHSNNITQVG VNDFCPIGFVKRAYNGISLFNNPVPYWEVQPA 350
               *****.*****.*****.*****

hBgn          351 TFRCVTDRLAIQFGNYRK 368
Xbgn AB037269 351 TFRCVTDRLAIQFGNYRK 368
               *****.*

```

B

```

Xbgn AB037269 253 ASLYRLGLGHNNIRMIENGSLSF MPVLRRELHLDNNKLSKVPPGLPDMK 300
Xbgn EST      1 ASLYRLGLGHNNIRMIENGSLSF MPVLRRELHLDNNKLSKVPPGLPDMK 48
               *****

Xbgn AB037269 301 LLQVVYLHSNNITQVG VNDFCPIGFVKRAYNGISLFNNPVPYWEVQPA 350
Xbgn EST      49 LLQVVYLHSNNITQVG VNDFCPIGFVKRAYNGISLFNNPVPYWEVQPA 98
               *****

Xbgn AB037269 351 TFRCVTDRLAIQFGNYRK 368
Xbgn EST      99 TFRCVTDRLAIQFGNYRK 116
               *****

```

Xbgn-EST and 5 nucleotide changes confined to a small region between nucleotides 500 (corresponding to base 1520 in AB037269) and 595 (1613 in AB037269) (See Appendix E). Closer analysis of the sequencing trace indicates that these changes are real and not an artefact of poor sequencing. Given the very close homology of *Xbgn-EST* to *Xbgn* AB037269 at the nucleotide level I believe these differences represent a small area of polymorphism, as opposed to indicating that *Xbgn-EST* is a pseudogene.

The temporal expression pattern of *Xbgn-EST* was determined by RT-PCR on total mRNA extracted from six different stages of embryonic development. *Xbgn-EST* transcripts are first detected at stage 13/14, followed by a large increase in expression between stage 35/36 and stage 42/43 (Figure 6.3A). Since *XBgn* has been shown to interact with collagen type 1 (Schonherr et al., 1995), we compared the expression of *Xbgn-EST* to that of *proColl1a1* and *proCol2a1*. *Xbgn-EST* transcripts are first detected before either collagen but the increase in expression levels between stage 35/36 and 42/43 also appears in these two genes. To determine the spatial expression pattern of *Xbgn-EST*, embryos were subjected to whole mount *in situ* hybridisation using an antisense probe for *Xbgn-EST*. Expression of *Xbgn-EST* was first detected by this technique at stage 16/17, in two small patches either side of the neural plate in cells thought to be neural crest (Figure 6.3Ba white arrows). Expression remains in this region at stage 24/25, either side of the neural plate (Figure 6.2Bb), but is more extensive and appears to be in the somites. At stage 33/34 expression levels have increased and transcripts are located to the notochord, the somites, posterior to the pharyngeal arches, and a small region in the extreme anterior of the head (Figure 6.3Bc).

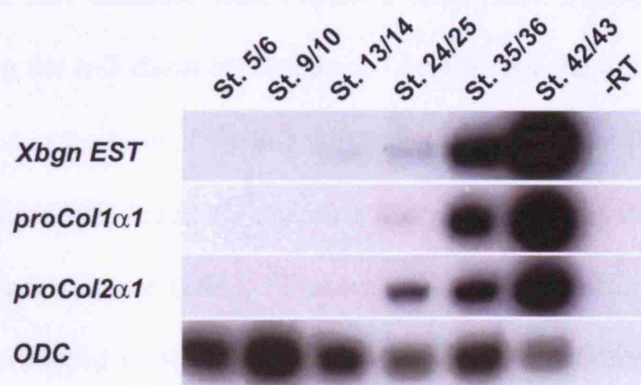
Figure 6.3 Expression of *Xbgn-EST*

A. The temporal expression pattern of *Xbgn-EST*, as determined by RT-PCR. Transcripts for *Xbgn-EST* were first detected at stage 13/14 and remained at similar levels at stage 24/25. Expression levels increased significantly between stage 24/25 and stage 35/36 and continued to increase through stage 42/43. For comparison, expression of *proColl1 α 1* and *proCol2 α 1* was also determined. Transcripts for both genes were first detected at stage 24/25. Expression levels for both genes increased in a similar fashion to that of *Xbgn-EST*. All transcript levels were compared to the uniformly expressed gene *ODC*. RNA from stage 24/25 embryos, minus reverse transcriptase (-RT), was used as a negative control. The result shown is representative of at least three independent experiments.

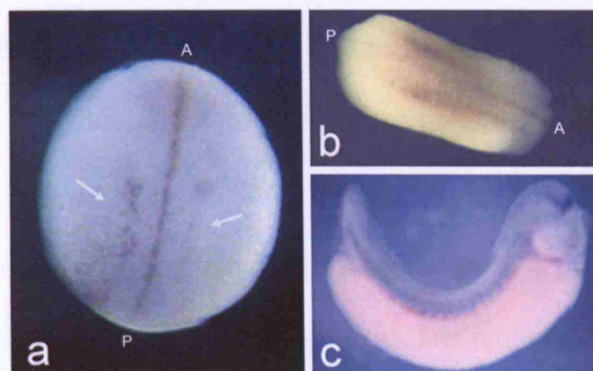
B. Expression of *Xbgn-EST* was also determined by whole mount *in situ* hybridisation using anti-sense DIG labelled probes. Expression is first detected by *in situ* at stage 13/14 in two patches either side of the middle section of the neural plate (a). Expression later appears in the somites (b) and at stage 35/36 transcripts are localised to the notochord, somites and specific regions of the head (c).

Figure 6.3

A



B



6.2.3 A laminin α -like chain is expressed in the notochord

BMP-1 cleaves the γ -2 chain of laminin-5 within the C-terminal LG domain of the protein and has been implicated in doing the same for the α -3 chain (Amano et al., 2000). The EST database search found a clone (Acc: BG892766) that was annotated as encoding the α -3 chain of laminin-5. An EST for the γ -2 chain was not present at the time and at the time of writing still remains absent from the databases. BG892766 was sequenced and found to contain a 1949 bp fragment that encoded a 367 amino acid sequence (Figure 6.4A). Performing a BLAST search with this amino acid sequence produced a closest match to a predicted chick laminin α -4 chain (E value $3.8e-72$), all other hits were for the laminin α -4 chain from other species. No *Xenopus* sequence, or α -3 chain, was returned in this search. Since BG892766 only appeared to contain the 3' end of the gene, we therefore turned to the EST database of *Xenopus tropicalis* (Sanger Institute) in an attempt to find a longer clone(s) that would help us with give a more accurate identification. Searching the *X. tropicalis* database with the *X. laevis* sequence identified a number of *X. tropicalis* clones that were assembled into a contiguous sequence of overlapping ESTs. This produced a sequence of 2677 bp, encoding a 528 amino acid fragment that was 88% identical to the *X. laevis* sequence in the overlapping region (Figure 6.4). A further BLAST search was performed using this longer sequence and returned a top hit of laminin α -5 chain from chick (E value $8.6e-98$), but was also closely related to laminin α -5 from humans. All α chains of laminin contain 5 G domains at their C-termini, which are involved in binding integrins and other ECM components such as syndecan and dystroglycan (Colognato and Yurchenco, 2000). Scanning the partial *X. laevis* protein sequence using the Prosite protein motif recognition program

Figure 6.4 Alignment of predicted protein sequence of *X. tropicalis* and *X. laevis* laminin α -like

The predicted laminin α -like protein sequence of *Xenopus tropicalis* was compared to the shorter sequence of *Xenopus laevis*. Laminin G domains (highlighted in grey) are approximately 170 amino acids in length and contain two conserved cysteines (blue letters). The *X. laevis* sequence consists of two G domains, while the *tropicalis* sequence contains an additional 127 amino acids of a third G domain and includes a site that follows the rules for a Tld cleavage site (red letters). There is also a conserved cysteine in the intervening sequence between first and second G domain of the *X. tropicalis* sequence (amino acid 167).

Figure 6.4

		G-Domain			
Xt Laminin alpha-like	1	RNEISTDRRPCWTAT*HDRLSVKILNAEMRSKTQFADGAGHYLTLYSNND	50		
Xl Laminin alpha-like	1		0		
Xt Laminin alpha-like	51	AVRFVDDKLOETRTLAAGTGRPDLTSTLLGGFQDGGGEIHNLTGCISNV	100		
Xl Laminin alpha-like	1		0		
Xt Laminin alpha-like	101	FVKRKDGPQAVLDLQQNQESLNVMTSCQASMEEQPQEIRSPSRKGLKQK	150		
Xl Laminin alpha-like	1		0		
		G-Domain			
Xt Laminin alpha-like	151	LRQKRLADHKHLEETSCVHPKAVRGALRFGGSSSSHLEFTRLPPIFRDRF	200		
Xl Laminin alpha-like	1	REDSSCVHPKAVKGALRFGGSSSSHLEFTRLPPIFRDRF	39		
		*,*****			
Xt Laminin alpha-like	201	HFSLEIRLNASNGLIFYGRDERSSSSLVSVNGRFVLQVLVGLKSLRLR	250		
Xl Laminin alpha-like	40	HFSLEIRLNTSNGLIFYGRDERSSSSLFLVSVNGRIVLHLVIGKSLRLR	89		

Xt Laminin alpha-like	251	SKEKYNDGLWHTVFFGKEKNKLHLVIDGIKAQSSVVYPVTSSTLTGPVFI	300		
Xl Laminin alpha-like	90	SKEKYNDGLWHTVFFGKEKNKLHLVIDGIKAQSSMVYPGVKSSLTGPVFI	139		

Xt Laminin alpha-like	301	GGLPPLIRRPDIPDASLNSFHGCLNKLKLDGKAFNPKNKIMGVAQYEGI	350		
Xl Laminin alpha-like	140	GGLPPLIRRPDIPDVSLSSFHGCLRNKLKLDGKALNPKNKIMGVTTQYEG	189		

		G-Domain			
Xt Laminin alpha-like	351	TEQGFFFSEGGFLKLDHEVDIGQSLEVKLEIRPIHHSGLLFHVGSEGGHS	400		
Xl Laminin alpha-like	190	TEQGFFFSEGGFLKLDHEVDIGQSLEVKLEIRPIHHSGLLFHVGTEEGHS	239		

Xt Laminin alpha-like	401	VTLSMAEGKVSVSVNAGASEYSTSVKFPQPLCDGQWHTVAVTKVSNVIQL	450		
Xl Laminin alpha-like	240	VTLSMTEGKVSVSVNAGASEYSTSVKFPQPLCDGQWHTVAVTKVSNVIQL	289		

Xt Laminin alpha-like	451	DVDTEGNHVGASQSQPTASTGTLVGGVPGNIHSPRKPHTPYQGCMRNL	500		
Xl Laminin alpha-like	290	DVDTEGNHVGASQTEPSRSGTLYIGGVPGNIYSPRKPSSYRGCMRNL	339		

Xt Laminin alpha-like	501	VINRKAVDVSKPGTFIGSAGTNVCPAL	528		
Xl Laminin alpha-like	340	VINRKAVDVSKPGTFVGSAGTNVCPSL	366		

(Sigrist et al., 2002) (<http://www.expasy.org/prosite>) revealed the presence of 2 laminin G domains between amino acids 15 & 185 and 191 & 363 (Figure 6.4 highlighted in grey). A similar scan using the longer *X. tropicalis* sequence included the last 127 amino acids of a third G domain and an intervening 50 amino acid sequence (Figure 6.4 highlighted in grey). Proteolytic cleavage of the γ -2 chain occurs between the third and fourth G domain, at a sequence that is conserved in humans, rat and mouse (Tsubota et al., 2000). This cleavage site was not present in the *X. tropicalis* sequence, but a putative cleavage site for a Tld protease was found within the 127 amino acid sequence of the third G-domain (Figure 6.4 amino acids 55 to 59, highlighted in red). The analysis performed here indicates that BG892766 encodes a laminin, but owing to the incomplete nature of the sequence it is impossible to determine which type of α -chain. I will therefore refer to it as laminin α -like. The nucleotide sequence of *laminin α -like* can be found in Appendix F.

The temporal expression pattern of *laminin α -like* was determined by RT-PCR on total RNA from six stages of *Xenopus* development (Figure 6.5A). *Laminin α -like* is a maternally expressed gene and is up-regulated at stage 9/10, the onset of gastrulation. Expression appears to remain at the same level throughout the stages tested, before a further increase in levels between stage 35/36 and stage 42/43. Spatial expression of *laminin α -like* was determined by whole mount *in situ* hybridisation. Transcripts are first detected in gastrulae and are specifically located to the region of presumptive notochord in dorsal marginal zone, extending anteriorly along the dorsal midline (Figure 6.5Ba). There was an apparent lengthening of the staining the further through gastrulation the embryo progressed. By stage 13/14 expression is localised to the notochord and an anterior region of the head (Figure

6.5Bb). At stage 33/34, expression is more widespread and is located to the notochord, otic vesicle, midbrain-hindbrain boundary (arrowhead), branchial arches, heart, pronephros and the pronephric duct (Figure 6.5Bd).

Figure 6.5 Expression of *laminin α -like*

A. Temporal expression of *laminin α -like* transcripts as determined by RT-PCR analysis of total RNA from six different stages of development. *Ornithine decarboxylase (ODC)* was used to control for different levels of cDNA in each sample. Stage 24/25 RNA minus reverse transcriptase (-RT) was used as a negative control. Maternal transcripts for *laminin α -like* were detected at stage 5/6. There is an increase in transcript levels at gastrulation (stage 9/10) and these remain relatively constant throughout the remaining stages tested. There does however appear to be a slight increase between stage 35/36 and stage 43/43. The result shown is representative of at least three independent experiments.

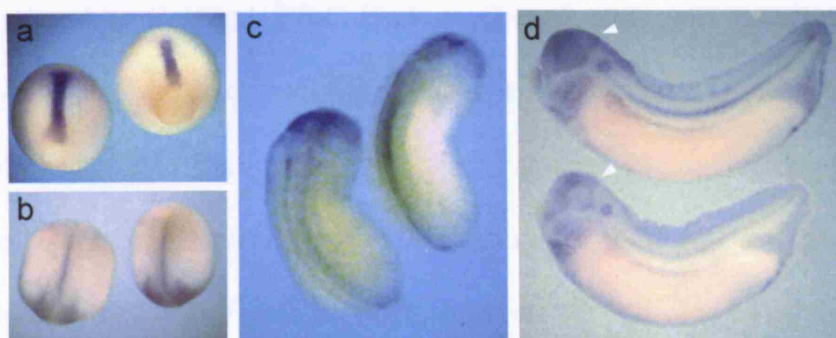
B. By *in situ* hybridisation, *laminin α -like* transcripts were first detected at gastrulation (stage 11/12) in the DMZ and presumptive notochord (a). Expression was later detected in the notochord proper and in the head at stage 17 (b) and stage 25 (c). At stage 33/34 expression was more widespread, with transcripts detected in the notochord, otic vesicle, pharyngeal arches, brain, heart, pronephros and pronephric duct.

Figure 6.5

A



B



6.3 Discussion

6.3.1 Fibrillar Collagens

The fibrillar collagens are a major group of proteins that provide structural stability in both the developing embryo and the adult. They are split into two groups, the major and minor, based on their abundance in adult tissues. Major collagens include type 1, type 2, and type 3, while minor collagens include type 5 and type 11 (Boot-Handford and Tuckwell, 2003). Both *proColl1 α* and *proCol2 α* encode major collagens that have previously been described in *Xenopus* (Goto et al., 2000; Su et al., 1991). In this chapter, I have identified four other members of this family present in *Xenopus* development; *proCol3 α* , *proCol5 α* , *proCol5 α 2*, and *proColl1 α* . *ProCol3 α* , like *proColl1 α* and *proCol2 α* is a major collagen fibre, which was first detected by RT-PCR at stage 24/25 and increased expression as development progressed. This is similar to what has been observed for the genes encoding the other major collagens, *proColl1 α* and *proColl1 α* . *ProCol5 α* , *proCol5 α 2*, and *proColl1 α* are all minor fibrillar collagens and like *proCol3 α* they were all detected by RT-PCR at stage 24/25, although *proCol5 α* is detected, faintly, at stage 13/14. Expression of these minor collagens is predominantly in the notochord, co-localising with transcripts for *proCol2 α* . *ProColl1 α* and *proCol2 α* transcripts are also found in the head, consistent with the fact that collagen type 11 is known to form heterotypic fibrils with collagen type 2, regulating collagen fibre diameter (Fernandes et al., 2003). Perhaps surprisingly, *proColl1 α* was not found in the somites, a major site of *proCol2 α* synthesis. Indeed, none of the minor collagens described in this chapter were expressed in the somites, even after long development times. The localisation of *proCol5 α* and *proCol5 α 2* to the notochord was unexpected, as collagen type 5 usually forms heterotypic fibrils with collagen type 1 and not collagen type 2 (Birk,

2001). We had expected these minor collagen genes to have a similar expression pattern to *proColl1 α 1*, which is localised to the dorsal epidermis (see Figure 6.1Ba) (Goto et al., 2000). One explanation for this observation is that type 2 and type 5 collagen can form heterotypic fibres in amphibians, all previous studies having been performed in mammals. Alternatively, these heterotypic fibres may be a feature of embryonic tissues, rather than adult tissues. Another plausible explanation is that type 5 collagen forms heterotypic fibres with other major (e.g. type 3) or minor (e.g. type 11) collagens (Mayne et al., 1993). Unfortunately I was unable to obtain a spatial expression pattern for *proCol3 α 1*, but in mammals type 3 collagen does associates with both type 1 and type 5 collagen (Boot-Handford and Tuckwell, 2003).

6.3.2 XBgn

Bgn is an SLRP and known substrate of BMP-1 in mammals (Scott et al., 2000). We obtained a *Xenopus laevis* EST (*Xbgn-EST*) displaying close homology to a known *Xenopus Bgn* sequence, in order to determine its expression pattern in developing embryos. *Xbgn-EST* contained a 1941 bp fragment that coded for the last 116 amino acids of *Xenopus Bgn*. An amino acid identity comparison indicates that the protein coded for by *Xbgn-EST* is identical to Xbgn, but with an incomplete 5' sequence it is impossible to be sure. In the coding region the nucleotide sequence of *Xbgn-EST* is identical to *Xbgn* AB037269, the only variation appears in the 3' UTR where there are 7 base pair changes. If *Xbgn-EST* were a pseudoallele one might expect one or two amino acid changes and/or a higher degree of variation in the nucleotide sequence. I am confident there are no sequencing errors in this region and therefore believe that the variation is more likely to be due to polymorphisms in this clone. Expression of *Xbgn-EST* was first detected at stage 13/14 by RT-PCR and transcripts level increased

significantly as development proceeded. The temporal pattern of expression mirrors that of *proColl α 1*, a protein with which Bgn is known to interact (Schonherr et al., 1995). However, the tissue distribution of these two genes is very different. While both genes are abundantly expressed in the dorsal half of the embryo, *proColl α 1* transcripts are localised to the inner layer of the epidermis and *Xbgn* transcripts to the somites (and later to the head and pronephros). This might reflect different functions for XBgn and type 1 collagen during *Xenopus* development.

Recent data has implicated XBgn in dorsal-ventral patterning of the marginal zone during gastrulation (Moreno et al., 2005), but this is not entirely consistent with the results presented in this chapter. Moreno et al. claim to have identified a *Xenopus* Bgn that is identical to the full-length sequence in the public database, which they refer to as AB037269, and therefore almost identical to the EST used in this study. Their results suggest that *Xbgn* is maternally expressed and present throughout the marginal zone during gastrulation. In addition, they show that it can induce secondary axes when mis-expressed in the VMZ. Although we have no functional studies with *Xbgn*, our expression data is not consistent with theirs. We find that *Xbgn* is not activated until after the onset of gastrulation and is localised to specific tissues such as the somites. We have no explanation for the differences of our results, since both sets of primers, ours and theirs, should detect the sequence in AB037269. One possible explanation is that we use different PCR conditions but this cannot explain the different *in situ* patterns. More detailed studies of *Xbgn* would be required to resolve this issue.

6.3.3 Laminin α -like

Laminin α -like was annotated in the EST database as a *laminin-5 α -3* chain (Acc: BG892766) and was of interest to us because mammalian laminin α -3 has been reported as a substrate of BMP-1 (Amano et al., 2000). Sequencing revealed the clone contained a 1.9kb C-terminal fragment encoding two G domain motifs that are found at the C-terminus of all laminin α -chains. Given the extremely large size of laminin genes (the longest human *laminin-5 α -3* transcript is 10.5 kb long (Accession number: NM198129)), a formal identity of this clone was not possible. The closest match in a database search with the laminin α -like amino acid sequence was the mammalian laminin α -4 chain. A similar search using a contiguous sequence of overlapping *X. tropicalis* ESTs displaying very close homology *X. laevis laminin α -like*, but encoding an extra 162 N-terminal amino acids, produced a closest match with mammalian laminin α -5. Since it is not clear whether this gene encodes a laminin α -4 chain or a laminin α -5, or indeed some other laminin α chain, we have called it *laminin α -like*.

As shown in Figure 6.4, the deduced *X. tropicalis* amino acid sequence for laminin α -like contains a putative BMP-1 cleavage site, in what we believe to be the third G-domain. Studies have suggested that proteolytic processing of the α -4 chain yields a fragment containing the last two LG domains, which play a role in interacting with different proteins in the ECM components (Talts et al., 2000; Yamashita et al., 2004). While there are no reports of proteolytic processing of laminin α -5, the C-terminal LG domains appear to play a similar role to that of Laminin α -4 by binding ECM proteins such as syndecans (Narita et al., 2004; Suzuki et al., 2003).

Early observations of laminin expression in *Xenopus* embryos showed that laminin β -1 was localised to the presumptive notochord and somitic mesoderm early in gastrulation (Fey and Hausen, 1990). We find that *laminin α -like* is maternally expressed and that it is localised to the presumptive notochord during gastrulation and to the notochord proper at later stages. Expression can also be found in the otic vesicle, brain, pronephros and pronephric duct. *Laminin α -like* expression in the dorsal midline of gastrulae was of particular interest as it suggests a potential role for this gene in dorsal gastrulation movements, such as convergent extension. Many reports suggest that laminin α chains have roles in the morphogenesis and disruption to either the formation or function of these chains disrupts gastrulation. In sea urchins, for example, injections of antibodies raised against two G-domains of the α -1 chain perturb gastrulation and embryos develop without or with very small archenterons (Benson et al., 1999). Genetic studies in *Drosophila* have shown that mutations in the genes for *laminin α -3* and *laminin α -5* disrupt morphogenesis (Garcia-Alonso et al., 1996; Henchcliffe et al., 1993) and in mice, mutations in the *laminin α -1* gene cause gastrulation defects (Miner et al., 2004). These studies indicate potential roles for *laminin α -like* in convergent extension. Indeed, the LG4 and LG5 domains of the laminin-5 α 3 chain promote cell migration (Tsubota et al., 2000) and presumptive *Xenopus* mesodermal cells will actively migrate on laminin coated cover-slips. suggesting an importance for this molecule in cell morphogenesis during gastrulation (Nakatsuji, 1986).

The localisation of the *laminin α -like* to the notochord is consistent with the apparent role for laminin proteins in notochord differentiation (Stemple, 2005). Two *zebrafish* mutants, *grumpy* and *sleepy*, display shortened body axes and defects in notochord

differentiation, and both genes encode laminins – *grumpy* encodes laminin β -1 and *sleepy* encodes laminin γ -2 (Parsons et al., 2002). There is a dramatic loss of laminin 1 staining in mutant embryos and the basement membrane that normally surrounds the notochord fails to form. Thus laminins play a key role in notochord differentiation. Further studies are necessary to define the role of laminin α -like in gastrulation and notochord differentiation, including injection of synthetic mRNA (over-expression) and antisense morpholinos (inhibition). However, the first step would be the isolation of cDNAs containing the entire ORF, a not inconsiderable problem since we only have ~2 kb of a potentially 10-11 kb mRNA.

6.3.4 Summary

This chapter identifies and describes the expression of a number of putative substrates for the Tlds expressed in early *Xenopus* development. *Xbgn* and *Col3 α 1*, *Col5 α 1*, *α 2* and *Col11a1* are all involved in collagen fibre formation and are expressed in the notochord, a structure surrounded by a sheath rich in collagen fibres. In addition, I have described the expression of a maternally expressed laminin that is localised to the presumptive notochord of gastrulae and later to the notochord proper, head, heart, and pronephros.

Chapter 7

General Discussion

7.1 The role of Tld metalloproteases

Members of the Tld family are closely related to the developmentally important *Drosophila* metalloprotease Tld. They have been found in both vertebrate and invertebrate species and are believed to play multiple roles in both embryonic development and adult physiology, roles that appear to be evolutionarily conserved. Two of these roles, in particular, have attracted the attention of embryologists; the control of dorsal-ventral patterning and assembly of the extracellular matrix. Studies in frogs, fish, and flies have concentrated on their role in dorsal-ventral patterning, while those in birds and mammals have concentrated on their role in the ECM.

In *Xenopus laevis*, three genes have been identified that encode members of the Tld family, BMP-1/Tld, Xld, and Xlr (Dale et al., 2002; Goodman et al., 1998; Lin et al., 1997; Maeno et al., 1993). These proteases cleave the BMP-4 binding protein Chd (Dale et al., 2002; Piccolo et al., 1997; Wardle et al., 1999), generating fragments with reduced affinity for BMP-4. As a consequence, Tld family members, when ectopically expressed, increase BMP-4 signalling and ventralise embryos (Goodman et al., 1998; Wardle et al., 1999). One of the aims of this thesis was to further characterise the role of Xld in regulating BMP-4 signalling, by identifying the domains required for the interaction with Chd. A second aim was to extend studies on the role of the Tlds in mammalian ECM formation into *Xenopus* development, by identifying *Xenopus laevis* homologues of mammalian ECM proteins processed by members of the Tld family.

7.2 The role of C-terminal CUB domains in Xld

A distinct problem for proteases with multiple substrates, all expressed in the same tissues or region of the embryo, is how they differentiate between different proteins with regards to cleavage. The Tlds are modular proteases consisting of an N-terminal protease domain and a C-terminal protein-protein interaction domain, composed of an array of CUB and EGF-like modules. In Chapter 4, I investigated the role of the CUB and EGF-like domains of Xld for recognising, and allowing proteolytic cleavage of, the dorsalising factor Chd. A previous study has shown that Xld cleaves Chd at two locations (Piccolo et al., 1997) but nothing is known about the regions of each protein that mediate the required protein-protein interaction. Xld contains five CUB and two EGF-like domains, yet I have shown that only the first and second CUB domains, along with the protease domain, are required to block the dorsalizing activity of Chd. Removing the third, fourth, and fifth CUB domains, as well as both EGF-like domains, has no effect on the ability of Xld to block the activity of Chd in the *Xenopus* embryo. The results are consistent with previous work on BMP-1, which showed that both the first and second CUB domains are required for efficient cleavage of proCol1 α 1 (Hartigan et al., 2003). However, whereas the second CUB domain is required for secretion of Xld, the first CUB domain suffices for secretion of mammalian BMP-1.

A number of future experiments need to be undertaken to extend this study. Primarily we need to show that Chd itself is physically cleaved by the various deletions in an *in vitro* assay. Our initial attempts using conditioned media from HEK293 cells were unsuccessful, both for full-length Xld and the Xld deletions. Since the media

contained secreted protease it is not immediately clear why full-length Xld did not cleave Chd, as we have previously had no difficulty in achieving this with Xld secreted from HEK293 cells (Piccolo et al., 1997) and *Xenopus* oocytes (Wardle et al., 1999). One explanation is that the Xld constructs used in this thesis added six-Myc-tags to the C-terminus, while previous constructs added either a single Flag-tag (Piccolo et al., 1997) or no tag at all (Wardle et al., 1999). Although Xld with six Myc-tags is active when expressed in *Xenopus* embryos (this thesis) it is possible that it is inactive in conditioned media, perhaps due to an unstable configuration. Xld may act soon after it has been secreted by injected embryonic cells, yet conditioned media was collected over 48 hours and frozen before processing. If the protease activity is unstable then this treatment might destroy the activity of Xld before use in cleavage assays. I have already remade all of the deletion constructs lacking these six Myc tags and will test their ability to cleave Chd *in vitro*.

A further series of experiments would be to test the activity of Xld, and the deletion constructs, against other putative substrates, such as the procollagens, XLox, and XBgn. Indeed, I had hoped to include such a study for XLox and/or XLoxl-1 in this thesis. Several attempts were made to put epitope tags onto the C-terminus of both of these proteins, as antibodies are not available for either protein, but this was unsuccessful. These are important experiments since they would tell us whether different CUB domains are used to target different substrates. For example, CUB3, 4, or 5 may be required for substrates other than Chd, thereby explaining their presence in Xld. However, if the 1st and 2nd CUB domains are sufficient for Xld to interact with, and cleave, all of its substrates, then we need to explain the role of the most C-terminal CUB domains. One explanation is that CUB3-5 are involved in directing the

protease to its correct location in the extracellular space, perhaps by interacting with the ECM. BMP-1 is known to bind heparin (Wozney et al., 1988) so this is not an implausible explanation.

If specific CUB domains are required for interactions between Xld and different groups of substrates, it should be possible to construct dominant-negative versions of the protease that specifically inhibit cleavage of these different substrates. This may allow us to confirm that the phenotype caused by expressing dnXld in *Xenopus* embryos (Piccolo et al., 1997; Wardle et al., 1999) is indeed the result of inhibiting Chd. Since the entire C-terminal half of Xld is included in this dominant-negative construct we cannot exclude a role for other substrates. Recent results have shown that both proCol2 α 1 and XBgn may have roles in regulating BMP-4 signalling during *Xenopus* development (Larrain et al., 2000; Moreno et al., 2005; Zhu et al., 1999) and cleavage of these proteins could contribute to the Xld over-expression phenotype. Protease minus constructs have been made for all of the Xld deletions and injected them into *Xenopus* embryos, but we found no effect on development. This was surprising, since even dnXld containing all of the CUB and EGF domains had no effect in this experiment, yet previous experiments have shown that it dorsalises ventral mesoderm (Wardle et al., 1999). Once again, six Myc-tags were added to the C-terminus of these “dominant-negatives” and I’m now concerned that they are having an adverse affect on either their secretion or stability. Constructs lacking these Myc-tags have now been made and will be tested for biological activity. If these new constructs are active they will be used in co-immunoprecipitation experiments to try and determine which domains physically bind Chd.

Finally, transcripts for the most recently described *Xenopus* Tld, *xlr*, are localised to ventral and lateral sectors of the marginal zone of gastrulae, a complementary expression pattern to that of *chd* (Dale et al., 2002). This suggests that Xlr is likely to play a major role in regulating the Chd/BMP-4 interaction in this region. With an identical structure to that of Xld, it would be interesting to determine if the same domains are involved in Xlr's role attenuating Chd, and if different can this shed more light on the role of the CUB domains in the Tlds.

7.3 Do inhibitors regulate Tld activity?

In *Xenopus*, XBMP-1, Xld and Xlr all counteract secondary axes induced by ectopic expression of *chd* mRNA in the VMZ (Dale et al., 2002; Piccolo et al., 1997; Wardle et al., 1999). It should stand to reason that they would do the same in the DMZ, where *chd* is expressed, yet both XBMP-1 and Xld are expressed in this region without adverse effect on the developing embryo. This raises the possibility that either they perform a different function in the organiser, or there is an endogenous inhibitor that protects Chd from cleavage within this region. Chapter 5 investigates the possibility that Edd, a close relative of the non-specific protease inhibitor α -2M (Sasai et al., 1996), might act in the capacity of a Tld inhibitor in the DMZ. Upon over-expression *edd* does not give rise to a noticeable phenotype, if it were a strong inhibitor of the Tlds one would expect it to dorsalise the embryo. Similarly, *edd* does not counteract the effect of co-injecting *xbmp-1* mRNA into the embryo, or its ability to counter act the axis inducing effect of *chd* in the secondary axis and VMZ assays, both more sensitive techniques to gross over-expression. α -2M binds proteases in a complex thereby precluding access of substrates to the enzyme. Under non-reducing conditions, Myc-tagged BMP-1 protein did not appear to increase in size, a sign that

other proteins were not bound and the presence of excess Edd did not change this. It was therefore clear that Edd was unlikely to be an inhibitor of BMP-1. In a final experiment, α -2M was added to media conditioned for Xlr, which alone was able to cleave Chd. The presence of α -2M was unable to affect the processing of Chd by Xlr and we therefore concluded that neither α -2M nor Edd are inhibitors of the Tlds.

The function of Edd in the embryo is still a mystery, although its close homology to α -2M suggests it is a protease inhibitor (Sasai et al., 1996). In this context, it is significant that Edd contains a conserved thio-ester group, which is important for binding proteases in the α -2M inhibitor complex (Sottrup-Jensen, 1989). However, to our knowledge, this thesis presents the first attempt to demonstrate a role for Edd as a protease inhibitor. Although Edd clearly does not inhibit Tld metalloproteases, it is possible that it inhibits other secreted proteases during *Xenopus* development. A number of secreted proteases have been described that are expressed during early *Xenopus* development, including members of the MMP (Harrison et al., 2004; Yang et al., 1997), ADAM (Alfandari et al., 2001; Alfandari et al., 1997; Smith et al., 2002), and serine protease (Yamada et al., 2000) families. Techniques similar to those used in this thesis could be applied to these proteases to test for inhibition by Edd, although this would be a long and tedious process if all *Xenopus* enzymes were to be tested. An alternative approach would be to further characterise the “bait-region” of Edd. The bait region of α -2M contains multiple cleavage sites for a wide variety of proteases (Sottrup-Jensen et al., 1989) and more detailed comparison of these sites and those in Edd might yield information as to the type of enzymes Edd may inhibit.

More specific inhibitors of metalloproteases have been described in mammals, including the tissue inhibitors of metalloproteases (TIMPs) that were first described as inhibitors of MMPs (Bode and Maskos, 2003). They therefore play important roles in regulating both ECM deposition and stability. However, TIMPS also inhibit some members of the ADAM family, which are involved in cell migration during embryogenesis (Alfandari et al., 2001; Smith et al., 2002). There is no published evidence that they inhibit members of the Tld family, but in co-injection experiments mammalian TIMP1 and TIMP2 did not block the ventralizing activity of XBMP-1 (Wardle and Dale, unpublished). Only TIMP-3 has so far been described in early *Xenopus* embryos, where it is both maternally and zygotically expressed (Pickard and Damjanovski, 2004; Yang and Kurkinen, 1998). Maternal transcripts are widely distributed, but from gastrula stages zygotic transcripts become enriched in dorsal structures. *XTIMP-3* RNA injected embryos display a range of different phenotypes, but most prominently head defects (Pickard and Damjanovski, 2004). The injected embryos were not analysed for defects in dorsal-ventral patterning that might be attributable to inhibition of Tlds. It would therefore be interesting to co-inject *Xenopus* embryos with *XTIMP-3* and Tld mRNAs, to determine whether the former is capable of inhibiting this family of metalloproteases.

7.3 Putative Tld substrates

Studies in mammals continue to identify new substrates for members of the Tld family and in particular for BMP-1. At least 16 different proteins have been reported as being processed by BMP-1 and it seems certain that this number will continue increase. It is also possible that substrates specifically cleaved by other members of the Tld family will be identified. However, few of the substrates identified in these

mammalian studies have so far been described in *Xenopus*, with most of the efforts in this organism directed at the role of these proteases in regulating Chd. Yet it is not inconceivable that some of these additional substrates contribute to the phenotypes obtained when Tlds are mis-expressed in *Xenopus* embryos, phenotypes that are currently solely ascribed to Chd (Dale et al., 2002; Goodman et al., 1998; Piccolo et al., 1997; Wardle et al., 1999). For example, proCol2 α 1 contains BMP-4 binding modules, similar to those of Chd, at its N-terminus and may have a role in regulating BMP signalling similar to that of Chd (Larrain et al., 2000; Sandell et al., 1991). Similarly, it has recently been suggested that XBgn has a role in regulating BMP signalling in *Xenopus* embryos (Moreno et al., 2005). Therefore, to fully understand the roles of Tlds in *Xenopus* development, it is essential that we know all of the substrates that are expressed at different stages of development.

In Chapters 3 and 6, I have identified and described the expression of a number of *Xenopus* genes encoding putative substrates for Tld metalloproteases, based on their homology to known substrates for mammalian Tlds. Initial identification came from EST database searches for *Xenopus* clones annotated as, or displaying significant homology to, proteins known to be BMP-1 substrates. In Chapter 3, I described the expression of three members of the Lox family. Our initial search for a Lox clone produced three different ESTs, all encoding proteins with the highly conserved LOX catalytic domain. Further sequencing revealed that these clones encoded three members of the family, two of which were full-length and most closely related to Lox and Loxl-1. In mammals both Lox and Loxl-1 are processed to a mature active form by BMP-1 (Borel et al., 2001; Uzel et al., 2001). Unfortunately, I was unable to confirm that this is also true of the *Xenopus* homologues. Several attempts to tag

these proteins at the C-terminus, with either Myc, His or Flag tags failed. It is not clear why this subcloning step should fail but I was unable to successfully transform *E. coli* with the tagged constructs. *Xenopus* Lox contains a conserved sequence that has been shown to be cleaved by BMP-1 in mammals (Panchenko et al., 1996; Uzel et al., 2001), so I anticipate that it will be cleaved by this protease. However, the BMP-1 cleavage site identified in mammalian Loxl-1 has not been conserved in *Xenopus* Loxl-1, so *in vitro* studies will be necessary to determine whether XLoxl-1 is a substrate for XBMP-1. Providing that a tagged version of the Lox proteins can be made, or specific antibodies generated, it would also be possible to determine whether they are cleaved by either Xld or Xlr.

Lox is involved in collagen fibre maturation, by allowing condensation of individual fibres into mature fibrils. It is therefore not surprising that *Xlox* transcripts are localised to the notochord, which is where most of the fibrillar collagens are also expressed. Collagen type II is a major component of the collagen sheath surrounding the notochord (Smith and Watt, 1985), so Lox, Loxl-1 and Loxl-3 are ideally located for enabling the formation of type II collagen fibres. The importance of Lox related enzymes to the notochordal sheath was shown in EM sections of embryos treated with the Lox specific inhibitor β -APN, in which collagen fibre formation was severely disrupted. Procollagen type I has also been described in *Xenopus* and transcripts are localised to the inner layer of the dorsal ectoderm (Goto et al., 2000). This tissue does not express any known Lox related enzymes and is not adjacent to the notochord, being separated by the somitic mesoderm, which probably precludes diffusion of Lox into this tissue. Either large collagen fibres are not formed in this tissue or other mechanisms are employed to promote their formation. This could involve a Lox

related enzyme that was not discovered in the screens employed in this thesis (e.g. Loxl-2 and Loxl-4). While I feel that I have investigated Lox function in the developing *Xenopus* embryo quite thoroughly, there are a couple of experiments that would be interesting to complete if we were to return to this research. For example, elastin is also a major substrate for Lox-like enzymes (Csiszar, 2001) and in mice Loxl-1 co-localises with elastin associated proteins in the ECM (Liu et al., 2004). I would therefore like to investigate the distribution of elastin in the early *Xenopus* embryo to see if this can shed further light on the role of the Lox-like enzymes. Unfortunately, *Xenopus* elastin has yet to be described. During *in silico* database searches I did find a *Xenopus* EST that was annotated as encoding *Xenopus* elastin, but sequencing demonstrated that this was an error. I failed to find elastin in searches of the *X. tropicalis* EST database (Sanger centre) and the *X. tropicalis* genome (JGI), using the human elastin mRNA sequence as a probe. The absence of a *Xenopus* EST is surprising, since elastin is an abundant protein in other organisms.

Of the major fibrillar collagen genes only *proColl1 α 1* and *proCol2 α 1* had been described in *Xenopus* prior to this study (Goto et al., 2000; Su et al., 1991). The third major member of this class, *proCol3 α 1*, had yet to be described. In addition to the major collagens, a number of minor accessory fibres exist, whose apparent role is to control and regulate the size of mature collagen fibrils predominantly made up of Col1, Col2 and Col3 (Fernandes et al., 2003; Garofalo et al., 1993; Wenstrup et al., 2004). Presumably different combinations of major and minor collagen give rise to fibres with different tensile strengths and properties depending on the site of expression. Again, no members of the minor group had been described in *Xenopus*. A search of the GenBank EST databases identified several ESTs that potentially

encoded many of these collagens and this was confirmed by sequencing. Transcripts for all three minor chains studied in this thesis (*proCol5 α 1*, α 2 and *proCol11 α 1*) localised to the notochord. The earliest expression of *proCol5 α 1*, *proCol5 α 2* and *proCol11 α 1* detected by RT-PCR was at stage 24/25, after initial expression of *proCol2 α 1*. Collagen type 5 is reported to control the fibre diameters of mature type 2 collagen and it was my expectation that at least one of the two chains to be detected as early as stage 13/14 when *proCol2 α 1* transcripts are first detected. However there is a large gap in developmental stages between stage 13/14 and stage 24/25 when the expression was analysed by RT-PCR. Ideally, a stage in between these two time points should have been included, although transcripts were detected by *in situ* hybridisation at much later stages indicating that, if expression is present it is found at very low levels. I have not determined the full amino-acid sequence for each of these collagen fibres, in part this was due to each EST being only a partial length and due to difficulties in sequencing the clones. The difficulties arose due to the high number of nucleotide repeats that code for the G-x-x amino acid repeats in the collagenous domain of each collagen. As a result sequencing becomes unreliable as the primers bind to multiple regions throughout the clone and prevents correct determination of the sequence. This problem could not be resolved in time to obtain a full sequence of each clone. It is unfortunate that the amino-acid sequence could not be obtained as this might give us an indication as to whether the BMP-1/PCP cleavage sites have been conserved in the *Xenopus* collagens.

Chapter 6 also describes the expression of *Xenopus* Bgn, which also plays a role in collagen fibre formation mammals by associating with collagen fibres and regulating their formation and diameter (Schonherr et al., 1995). *Xenopus* Bgn contains a short

amino acid sequence near its N-terminus that is nearly identical to the sequence in mammalian Bgn that has been shown to be cleaved by BMP-1 (Scott et al., 2000). Thus XBgn is probably cleaved by XBMP-1, although I have not formally shown this to be the case. Recently, XBgn has been shown to induce secondary axes in *Xenopus* embryos, presumably because it inhibits BMP signalling (Moreno et al., 2005). However, there is a discrepancy between this data and the data presented in this thesis. While Moreno et al. detect *Xbgn* as a maternal transcript, and in the marginal zone at gastrulation, my results indicate expression is much later and similar to that of the fibrillar collagen genes *proColl α 1* and *proCol2 α 1*. The nucleotide sequence of *XBgn-EST* is identical to the full-length *XBgn* found in the EST database, which Moreno et al. claim their *XBgn* clone is identical to. The PCR primers used in this study and those used by Moreno et al. should identify the same gene, thereby giving identical results. In an attempt to resolve this issue I have obtained the PCR primers of Moreno et al. and I will use them to try and repeat their results. I was unable to do this before writing up this thesis.

A sixth EST was identified because it was annotated as encoding the laminin α -3 chain, and thought to be a substrate of BMP-1 in mammals (Amano et al., 2000). However, upon sequencing it was found to be a small 2 kb C-terminal fragment of a much larger laminin clone that did not appear to be an α -3 chain as described. Using *in silico* techniques, a contig of homologous *X. tropicalis* sequences was constructed in an attempt to obtain a more accurate identity for the clone. The predicted amino acid sequence displayed closest homology to both laminin α -4 and laminin α -5, and since a formal identification could not be made we have named this clone *laminin α -like*. A common characteristic of all laminin α -chains is the presence of C-terminal

globular domains (LG), which mediate laminin macromolecules' interaction with the basal lamina and other components of the ECM, often via proteolytic processing (Colognato and Yurchenco, 2000). These LG domains have been reported to mediate cell migration *in vitro* (Tsubota et al., 2000; Tsubota et al., 2005) and antibodies raised against them will disrupt gastrulation in sea urchin embryo (Benson et al., 1999), suggesting a role for laminin chains in the movements of convergent-extension. Consistent with this I found *laminin α -like* expressed in the dorsal midline of gastrulating embryos, an area undergoing many convergent extension movements (Keller et al., 2003). This region gives rise to the notochord, which expresses laminins (Fey and Hausen, 1990), and their knock out disrupts notochord formation (Parsons et al., 2002). While it is unlikely that I will clone the full-length laminin, for which we have only the 3' end, an interesting experiment could involve over-expressing the two LG domains this clone codes for in early *Xenopus* embryos. By adding a signal peptide from an extracellular protein such as activin to the 5' end of this construct, it should be possible to get these domains secreted to the ECM. The results of Benson et al. (1999) show that an antibody binding the C-terminal LG domains of laminin α -1 disrupts sea urchin gastrulation and it would be interesting to see if the over-expression of the LG domains of laminin α -like has a similar effect.

7.4 Closing thoughts

This thesis brings together two aspects of Tld activity in *Xenopus* embryos, in an attempt to extend our knowledge of the function and control of XBMP-1/Tld, Xld and Xlr. It is now apparent that many substrates exist for the Tld family and this thesis has identified the presence in *Xenopus* embryos of eight genes that code for proteins potentially processed by these proteases. The ability to differentiate between multiple

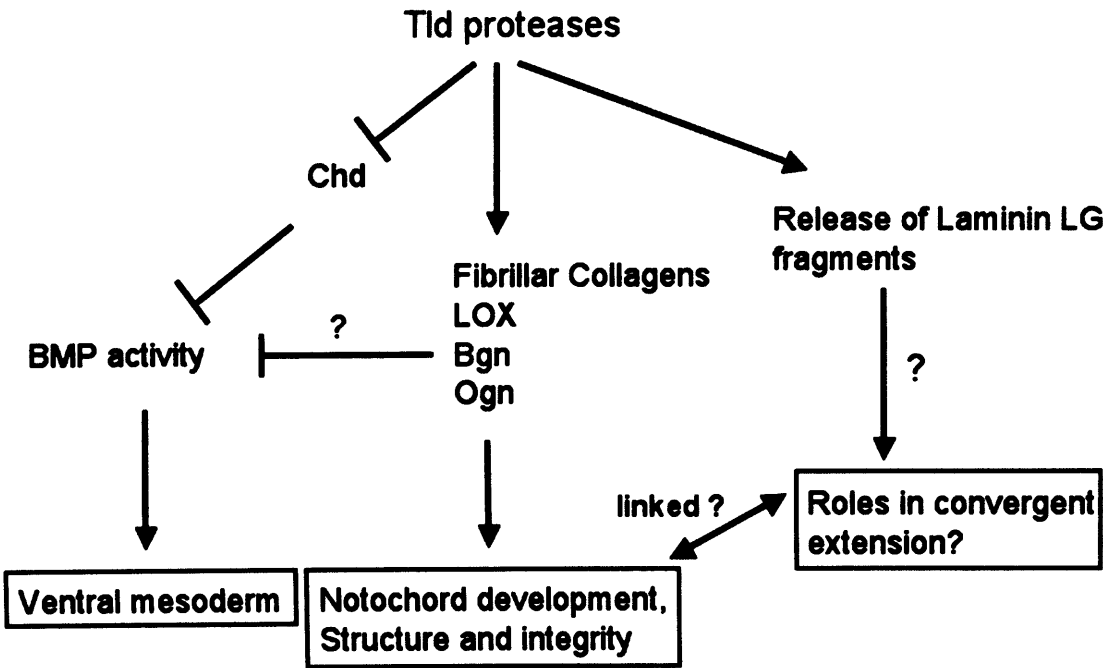
substrates is a problem for any enzyme and we are just beginning to understand how each module of the Tlds functions to facilitate this. Although we have insufficient data to draw a model of how the Tlds interact with various substrates, one can envisage a scenario whereby the presence and availability of different domains in the Tlds differentiate between substrates in the ECM. However, another scenario is that there is little differentiation between the substrates and differences in their cleavage are simply a result of their temporal-spatial distribution.

While our understanding of the role Tlds play in mesoderm patterning and gastrulation is still in its infancy, I have started to gain insights into how the various components they control might interact with one another in the developing *Xenopus* embryo. The Tlds help pattern the marginal zone by regulating the interaction of Chd with BMP-4, which, if the report of Moreno et al., (2005) is correct, is also influenced by XBgn. In addition, they control the formation of collagen fibres in the ECM, in particular the notochord, which arises from the DMZ and stiffens the embryo during elongation. A further level of control over signalling in the DMZ and marginal zone might come from the availability of the pro-peptides yielded after cleavage by the Tlds, as reports are starting to emerge of pro-peptides wielding their own biological activity, including those from proCol2 α 1 and Lox (Larrain et al., 2000; Palamakumbura et al., 2004). In regulating fragments of laminin chains, which stimulate cellular movements, the Tlds may also have a level of control over convergent-extension movements in the gastrulating embryo. A diagram summarising these thoughts is presented in Figure 7.1.

Figure 7.1. A diagram summarising my current thoughts on Tld function in early embryonic development.

Members of the Tld family prevent Chd from inhibiting BMP-4 therefore promoting the formation of ventral cell fates. In addition, they activate many members of the fibrillar collagen family and other proteins necessary for correct collagen formation, of these Bgn is may also have an inhibitory effect on BMP activity. The fibrillar collagens play important roles in the formation and maintenance of the notochord, the precursors of which are laid down on the dorsal side of the embryo during the convergent extension movements of gastrulation. LG domains of certain laminin subunits, thought to be released via proteolytic cleavage by members of the Tld family, may also play a role in convergent extension movements in the DMZ and formation of the notochord.

Figure 7.1



References

- Adams, D. S., Keller, R. and Koehl, M. A.** (1990). The mechanics of notochord elongation, straightening and stiffening in the embryo of *Xenopus laevis*. *Development* **110**, 115-30.
- Adar, R., Kessler, E. and Goldberg, B.** (1986). Evidence for a protein that enhances the activity of type I procollagen C-proteinase. *Coll Relat Res* **6**, 267-77.
- Agius, E., Oelgeschlager, M., Wessely, O., Kemp, C. and De Robertis, E. M.** (2000). Endodermal Nodal-related signals and mesoderm induction in *Xenopus*. *Development* **127**, 1173-83.
- Ahmed, N., Howard, L. and Woodland, H. R.** (2004). Early endodermal expression of the *Xenopus* Endodermin gene is driven by regulatory sequences containing essential Sox protein-binding elements. *Differentiation* **72**, 171-84.
- Alfandari, D., Cousin, H., Gaultier, A., Smith, K., White, J. M., Darribere, T. and DeSimone, D. W.** (2001). *Xenopus* ADAM 13 is a metalloprotease required for cranial neural crest-cell migration. *Curr Biol* **11**, 918-30.
- Alfandari, D., Wolfsberg, T. G., White, J. M. and DeSimone, D. W.** (1997). ADAM 13: a novel ADAM expressed in somitic mesoderm and neural crest cells during *Xenopus laevis* development. *Dev Biol* **182**, 314-30.
- Amano, S., Scott, I. C., Takahara, K., Koch, M., Champliand, M. F., Gerecke, D. R., Keene, D. R., Hudson, D. L., Nishiyama, T., Lee, S. et al.** (2000). Bone morphogenetic protein 1 is an extracellular processing enzyme of the laminin 5 gamma 2 chain. *J Biol Chem* **275**, 22728-35.
- Amaya, E., Musci, T. J. and Kirschner, M. W.** (1991). Expression of a dominant negative mutant of the FGF receptor disrupts mesoderm formation in *Xenopus* embryos. *Cell* **66**, 257-70.
- Arandjelovic, S., Freed, T. A. and Gonias, S. L.** (2003). Growth factor-binding sequence in human alpha2-macroglobulin targets the receptor-binding site in transforming growth factor-beta. *Biochemistry* **42**, 6121-7.
- Asashima, M., Nakano, H., Shimada, K., Kinishoto, K., Ishii, K., Shibai, H. and Ueno, N.** (1990). Mesodermal induction in early amphibian embryos by activin A (erythroid differentiation factor). *Wilhelm Roux's Arch Dev Biol* **198**, 330-335.
- Asuncion, L., Fogelgren, B., Fong, K. S., Fong, S. F., Kim, Y. and Csiszar, K.** (2001). A novel human lysyl oxidase-like gene (LOXL4) on chromosome 10q24 has an altered scavenger receptor cysteine rich domain. *Matrix Biol* **20**, 487-91.
- Barrett, A. J.** (1981). Alpha 2-macroglobulin. *Methods Enzymol* **80 Pt C**, 737-54.
- Benson, S., Page, L., Ingersoll, E., Rosenthal, E., Dungca, K. and Signor, D.** (1999). Developmental characterization of the gene for laminin alpha-chain in sea urchin embryos. *Mech Dev* **81**, 37-49.
- Bieker, J. J. and Yazdani-Buicky, M.** (1992). Distribution of type II collagen mRNA in *Xenopus* embryos visualized by whole-mount in situ hybridization. *J Histochem Cytochem* **40**, 1117-20.
- Birk, D. E.** (2001). Type V collagen: heterotypic type I/V collagen interactions in the regulation of fibril assembly. *Micron* **32**, 223-37.
- Birsoy, B., Berg, L., Williams, P. H., Smith, J. C., Wylie, C. C., Christian, J. L. and Heasman, J.** (2005). XPACE4 is a localized pro-protein convertase required for mesoderm induction and the cleavage of specific TGFbeta proteins in *Xenopus* development. *Development* **132**, 591-602.

- Blader, P., Rastegar, S., Fischer, N. and Strahle, U.** (1997). Cleavage of the BMP-4 antagonist chordin by zebrafish tolloid. *Science* **278**, 1937-40.
- Blitz, I. L., Shimmi, O., Wunnenberg-Stapleton, K., O'Connor, M. B. and Cho, K. W.** (2000). Is chordin a long-range- or short-range-acting factor? Roles for BMP1-related metalloproteases in chordin and BMP4 autofeedback loop regulation. *Dev Biol* **223**, 120-38.
- Blumberg, B., Wright, C. V., De Robertis, E. M. and Cho, K. W.** (1991). Organizer-specific homeobox genes in *Xenopus laevis* embryos. *Science* **253**, 194-6.
- Bode, W., Gomis-Ruth, F. X., Huber, R., Zwillig, R. and Stocker, W.** (1992). Structure of astacin and implications for activation of astacins and zinc-ligation of collagenases. *Nature* **358**, 164-7.
- Bode, W., Grams, F., Reinemer, P., Gomis-Ruth, F. X., Baumann, U., McKay, D. B. and Stocker, W.** (1994). The metzincins: A new superfamily of Zinc endopeptidases. In *The Astacins, Structure and Function of a New Protein Family*, (ed. R. Zwillig and W. Stocker). Hamburg: Verlag Dr. Kovac.
- Bode, W. and Maskos, K.** (2003). Structural basis of the matrix metalloproteinases and their physiological inhibitors, the tissue inhibitors of metalloproteinases. *Biol Chem* **384**, 863-72.
- Bond, J. S. and Beynon, R. J.** (1995). The astacin family of metalloendopeptidases. *Protein Sci* **4**, 1247-61.
- Boot-Handford, R. P. and Tuckwell, D. S.** (2003). Fibrillar collagen: the key to vertebrate evolution? A tale of molecular incest. *Bioessays* **25**, 142-51.
- Boot-Handford, R. P., Tuckwell, D. S., Plumb, D. A., Rock, C. F. and Poulson, R.** (2003). A novel and highly conserved collagen (pro(α 1)(XXVII)) with a unique expression pattern and unusual molecular characteristics establishes a new clade within the vertebrate fibrillar collagen family. *J Biol Chem* **278**, 31067-77.
- Borel, A., Eichenberger, D., Farjanel, J., Kessler, E., Gleyzal, C., Hulmes, D. J., Sommer, P. and Font, B.** (2001). Lysyl oxidase-like protein from bovine aorta. Isolation and maturation to an active form by bone morphogenetic protein-1. *J Biol Chem* **276**, 48944-9.
- Bork, P.** (1991). Complement components C1r/C1s, bone morphogenic protein 1 and *Xenopus laevis* developmentally regulated protein UVS.2 share common repeats. *FEBS Lett* **282**, 9-12.
- Bork, P. and Beckmann, G.** (1993). The CUB domain. A widespread module in developmentally regulated proteins. *J Mol Biol* **231**, 539-45.
- Bouwmeester, T., Kim, S., Sasai, Y., Lu, B. and De Robertis, E. M.** (1996). Cerberus is a head-inducing secreted factor expressed in the anterior endoderm of Spemann's organizer. *Nature* **382**, 595-601.
- Busby, T. F. and Ingham, K. C.** (1988). Domain structure, stability, and interactions of human complement C1s: characterization of a derivative lacking most of the B chain. *Biochemistry* **27**, 6127-35.
- Butler, E., Hardin, J. and Benson, S.** (1987). The role of lysyl oxidase and collagen crosslinking during sea urchin development. *Exp Cell Res* **173**, 174-82.
- Chalmers, A. D. and Slack, J. M.** (1998). Development of the gut in *Xenopus laevis*. *Dev Dyn* **212**, 509-21.
- Chang, C., Holtzman, D. A., Chau, S., Chickering, T., Woolf, E. A., Holmgren, L. M., Bodorova, J., Gearing, D. P., Holmes, W. E. and Brivanlou, A. H.** (2001). Twisted gastrulation can function as a BMP antagonist. *Nature* **410**, 483-7.

- Childs, S. R. and O'Connor, M. B.** (1994). Two domains of the tolloid protein contribute to its unusual genetic interaction with decapentaplegic. *Dev Biol* **162**, 209-20.
- Clements, D., Cameleyre, I. and Woodland, H. R.** (2003). Redundant early and overlapping larval roles of Xsox17 subgroup genes in *Xenopus* endoderm development. *Mech Dev* **120**, 337-48.
- Colognato, H. and Yurchenco, P. D.** (2000). Form and function: the laminin family of heterotrimers. *Dev Dyn* **218**, 213-34.
- Conlon, F. L., Lyons, K. M., Takaesu, N., Barth, K. S., Kispert, A., Herrmann, B. and Robertson, E. J.** (1994). A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. *Development* **120**, 1919-28.
- Csiszar, K.** (2001). Lysyl oxidases: a novel multifunctional amine oxidase family. *Prog Nucleic Acid Res Mol Biol* **70**, 1-32.
- Dale, L., Evans, W. and Goodman, S. A.** (2002). Xolloid-related: a novel BMP1/Tolloid-related metalloprotease is expressed during early *Xenopus* development. *Mech Dev* **119**, 177-90.
- Dale, L., Howes, G., Price, B. M. and Smith, J. C.** (1992). Bone morphogenetic protein 4: a ventralizing factor in early *Xenopus* development. *Development* **115**, 573-85.
- Dale, L. and Jones, C. M.** (1999). BMP signalling in early *Xenopus* development. *Bioessays* **21**, 751-60.
- Dale, L., Matthews, G. and Colman, A.** (1993). Secretion and mesoderm-inducing activity of the TGF-beta-related domain of *Xenopus* Vg1. *Embo J* **12**, 4471-80.
- Dale, L. and Slack, J. M.** (1987). Regional specification within the mesoderm of early embryos of *Xenopus laevis*. *Development* **100**, 279-95.
- Dale, L., Smith, J. C. and Slack, J. M.** (1985). Mesoderm induction in *Xenopus laevis*: a quantitative study using a cell lineage label and tissue-specific antibodies. *J Embryol Exp Morphol* **89**, 289-312.
- Dale, L. and Wardle, F. C.** (1999). A gradient of BMP activity specifies dorsal-ventral fates in early *Xenopus* embryos. *Semin Cell Dev Biol* **10**, 319-26.
- Danielpour, D. and Sporn, M. B.** (1990). Differential inhibition of transforming growth factor beta 1 and beta 2 activity by alpha 2-macroglobulin. *J Biol Chem* **265**, 6973-7.
- Dawson, D. A.** (1993). Joint action of benzoic hydrazide and beta-aminopropionitrile on *Xenopus* embryo development. *Toxicology* **81**, 123-30.
- Dawson, D. A., Cotter, M. A., Policz, D. L., Stoffer, D. A., Nichols, J. P. and Poch, G.** (2000). Comparative evaluation of the combined osteolathyrin effects of two nitrile combinations on *xenopus* embryos. *Toxicology* **147**, 193-207.
- Dawson, D. A., Rinaldi, A. C. and Poch, G.** (2002). Biochemical and toxicological evaluation of agent-cofactor reactivity as a mechanism of action for osteolathyrism. *Toxicology* **177**, 267-84.
- Dawson, D. A., Schultz, T. W., Baker, L. L. and Mannar, A.** (1990). Structure-activity relationships for osteolathyrism. III. Substituted thiosemicarbazides. *J Appl Toxicol* **10**, 59-64.
- De Robertis, E. M. and Kuroda, H.** (2004). Dorsal-ventral patterning and neural induction in *Xenopus* embryos. *Annu Rev Cell Dev Biol* **20**, 285-308.
- De Robertis, E. M. and Sasai, Y.** (1996). A common plan for dorsoventral patterning in Bilateria. *Nature* **380**, 37-40.
- Delgadillo-Reynoso, M. G., Rollo, D. R., Hursh, D. A. and Raff, R. A.** (1989). Structural analysis of the uEGF gene in the sea urchin *strongylocentrotus purpuratus*

reveals more similarity to vertebrate than to invertebrate genes with EGF-like repeats. *J Mol Evol* **29**, 314-27.

Dennis, P. A., Saksela, O., Harpel, P. and Rifkin, D. B. (1989). Alpha 2-macroglobulin is a binding protein for basic fibroblast growth factor. *J Biol Chem* **264**, 7210-6.

Dosch, R., Gawantka, V., Delius, H., Blumenstock, C. and Niehrs, C. (1997). Bmp-4 acts as a morphogen in dorsoventral mesoderm patterning in *Xenopus*. *Development* **124**, 2325-34.

Dumermuth, E., Sterchi, E. E., Jiang, W. P., Wolz, R. L., Bond, J. S., Flannery, A. V. and Beynon, R. J. (1991). The astacin family of metalloendopeptidases. *J Biol Chem* **266**, 21381-5.

Elkan, E. R. (1938). The *Xenopus* pregnancy test. *Br Med J* **2**, 1253-1256.

Fainsod, A., Deissler, K., Yelin, R., Marom, K., Epstein, M., Pillemer, G., Steinbeisser, H. and Blum, M. (1997). The dorsalizing and neural inducing gene follistatin is an antagonist of BMP-4. *Mech Dev* **63**, 39-50.

Fan, T. J. and Katagiri, C. (2001). Properties of the hatching enzyme from *Xenopus laevis*. *Eur J Biochem* **268**, 4892-8.

Feinberg, H., Uitdehaag, J. C., Davies, J. M., Wallis, R., Drickamer, K. and Weis, W. I. (2003). Crystal structure of the CUB1-EGF-CUB2 region of mannose-binding protein associated serine protease-2. *Embo J* **22**, 2348-59.

Ferguson, E. L. and Anderson, K. V. (1991). Dorsal-ventral pattern formation in the *Drosophila* embryo: the role of zygotically active genes. *Curr Top Dev Biol* **25**, 17-43.

Ferguson, E. L. and Anderson, K. V. (1992a). Decapentaplegic acts as a morphogen to organize dorsal-ventral pattern in the *Drosophila* embryo. *Cell* **71**, 451-61.

Ferguson, E. L. and Anderson, K. V. (1992b). Localized enhancement and repression of the activity of the TGF-beta family member, decapentaplegic, is necessary for dorsal-ventral pattern formation in the *Drosophila* embryo. *Development* **114**, 583-97.

Fernandes, R. J., Schmid, T. M. and Eyre, D. R. (2003). Assembly of collagen types II, IX and XI into nascent hetero-fibrils by a rat chondrocyte cell line. *Eur J Biochem* **270**, 3243-50.

Fey, J. and Hausen, P. (1990). Appearance and distribution of laminin during development of *Xenopus laevis*. *Differentiation* **42**, 144-52.

Finelli, A. L., Bossie, C. A., Xie, T. and Padgett, R. W. (1994). Mutational analysis of the *Drosophila* tolloid gene, a human BMP-1 homolog. *Development* **120**, 861-70.

Finelli, A. L., Xie, T., Bossie, C. A., Blackman, R. K. and Padgett, R. W. (1995). The tolkin gene is a tolloid/BMP-1 homologue that is essential for *Drosophila* development. *Genetics* **141**, 271-81.

Francois, V. and Bier, E. (1995). *Xenopus* chordin and *Drosophila* short gastrulation genes encode homologous proteins functioning in dorsal-ventral axis formation. *Cell* **80**, 19-20.

Francois, V., Solloway, M., O'Neill, J. W., Emery, J. and Bier, E. (1994). Dorsal-ventral patterning of the *Drosophila* embryo depends on a putative negative growth factor encoded by the short gastrulation gene. *Genes Dev* **8**, 2602-16.

Gacheru, S. N., Trackman, P. C., Shah, M. A., O'Gara, C. Y., Spacciapoli, P., Greenaway, F. T. and Kagan, H. M. (1990). Structural and catalytic properties of copper in lysyl oxidase. *J Biol Chem* **265**, 19022-7.

- Garcia-Alonso, L., Fetter, R. D. and Goodman, C. S.** (1996). Genetic analysis of Laminin A in *Drosophila*: extracellular matrix containing laminin A is required for ocellar axon pathfinding. *Development* **122**, 2611-21.
- Garofalo, S., Metsaranta, M., Ellard, J., Smith, C., Horton, W., Vuorio, E. and de Crombrughe, B.** (1993). Assembly of cartilage collagen fibrils is disrupted by overexpression of normal type II collagen in transgenic mice. *Proc Natl Acad Sci U S A* **90**, 3825-9.
- Garrigue-Antar, L., Francois, V. and Kadler, K. E.** (2004). Deletion of epidermal growth factor-like domains converts mammalian tolloid into a chordinase and effective procollagen C-proteinase. *J Biol Chem* **279**, 49835-41.
- Ge, G., Seo, N. S., Liang, X., Hopkins, D. R., Hook, M. and Greenspan, D. S.** (2004). Bone morphogenetic protein-1/tolloid-related metalloproteinases process osteoglycin and enhance its ability to regulate collagen fibrillogenesis. *J Biol Chem* **279**, 41626-33.
- Geach, T. J. and Dale, L.** (2005). Members of the lysyl oxidase family are expressed during development of the frog *Xenopus laevis*. *Differentiation* **In press**.
- Geier, G. and Zwilling, R.** (1998). Cloning and characterization of a cDNA coding for *Astacus* embryonic astacin, a member of the astacin family of metalloproteases from the crayfish *Astacus astacus*. *Eur J Biochem* **253**, 796-803.
- Gerhart, J., Danilchik, M., Doniach, T., Roberts, S., Rowning, B. and Stewart, R.** (1989). Cortical rotation of the *Xenopus* egg: consequences for the anteroposterior pattern of embryonic dorsal development. *Development* **107 Suppl**, 37-51.
- Giampuzzi, M., Botti, G., Di Duca, M., Arata, L., Ghiggeri, G., Gusmano, R., Ravazzolo, R. and Di Donato, A.** (2000). Lysyl oxidase activates the transcription activity of human collagen III promoter. Possible involvement of Ku antigen. *J Biol Chem* **275**, 36341-9.
- Giampuzzi, M., Oleggini, R. and Di Donato, A.** (2003). Altered adhesion features and signal transduction in NRK-49F cells transformed by down-regulation of lysyl oxidase. *Biochim Biophys Acta* **1647**, 239-44.
- Giger, R. J., Urquhart, E. R., Gillespie, S. K., Levengood, D. V., Ginty, D. D. and Kolodkin, A. L.** (1998). Neuropilin-2 is a receptor for semaphorin IV: insight into the structural basis of receptor function and specificity. *Neuron* **21**, 1079-92.
- Godin, R. E., Klinzing, D. C., Porcaro, W. A. and Ernst, S. G.** (1997). Specification of endoderm in the sea urchin embryo. *Mech Dev* **67**, 35-47.
- Gomis-Ruth, F. X., Stocker, W., Huber, R., Zwilling, R. and Bode, W.** (1993). Refined 1.8 Å X-ray crystal structure of astacin, a zinc-endopeptidase from the crayfish *Astacus astacus* L. Structure determination, refinement, molecular structure and comparison with thermolysin. *J Mol Biol* **229**, 945-68.
- Gonzalez, E. M., Reed, C. C., Bix, G., Fu, J., Zhang, Y., Gopalakrishnan, B., Greenspan, D. S. and Iozzo, R. V.** (2005). BMP-1/Tolloid-like metalloproteases process endorepellin, the angiostatic C-terminal fragment of perlecan. *J Biol Chem* **280**, 7080-7.
- Goodman, S. A., Albano, R., Wardle, F. C., Matthews, G., Tannahill, D. and Dale, L.** (1998). BMP1-related metalloproteinases promote the development of ventral mesoderm in early *Xenopus* embryos. *Dev Biol* **195**, 144-57.
- Goto, T., Katada, T., Kinoshita, T. and Kubota, H. Y.** (2000). Expression and characterization of *Xenopus* type I collagen alpha 1 (COL1A1) during embryonic development. *Dev Growth Differ* **42**, 249-56.

- Graff, J. M., Thies, R. S., Song, J. J., Celeste, A. J. and Melton, D. A.** (1994). Studies with a *Xenopus* BMP receptor suggest that ventral mesoderm- inducing signals override dorsal signals in vivo. *Cell* **79**, 169-79.
- Green, J. B. and Smith, J. C.** (1990). Graded changes in dose of a *Xenopus* activin A homologue elicit stepwise transitions in embryonic cell fate. *Nature* **347**, 391-4.
- Grinnell, F., Zhu, M. and Parks, W. C.** (1998). Collagenase-1 complexes with alpha2-macroglobulin in the acute and chronic wound environments. *J Invest Dermatol* **110**, 771-6.
- Gurdon, J. B. and Hopwood, N.** (2000). The introduction of *Xenopus laevis* into developmental biology: of empire, pregnancy testing and ribosomal genes. *Int J Dev Biol* **44**, 43-50.
- Hamalainen, E. R., Jones, T. A., Sheer, D., Taskinen, K., Pihlajaniemi, T. and Kivirikko, K. I.** (1991). Molecular cloning of human lysyl oxidase and assignment of the gene to chromosome 5q23.3-31.2. *Genomics* **11**, 508-16.
- Handford, P. A., Baron, M., Mayhew, M., Willis, A., Beesley, T., Brownlee, G. G. and Campbell, I. D.** (1990). The first EGF-like domain from human factor IX contains a high-affinity calcium binding site. *Embo J* **9**, 475-80.
- Haramoto, Y., Tanegashima, K., Onuma, Y., Takahashi, S., Sekizaki, H. and Asashima, M.** (2004). *Xenopus tropicalis* nodal-related gene 3 regulates BMP signaling: an essential role for the pro-region. *Dev Biol* **265**, 155-68.
- Harland, R. and Gerhart, J.** (1997). Formation and function of Spemann's organizer. *Annu Rev Cell Dev Biol* **13**, 611-67.
- Harland, R. M.** (1991). In situ hybridization: an improved whole-mount method for *Xenopus* embryos. *Methods Cell Biol* **36**, 685-95.
- Harrison, M., Abu-Elmagd, M., Grocott, T., Yates, C., Gavrilovic, J. and Wheeler, G. N.** (2004). Matrix metalloproteinase genes in *Xenopus* development. *Dev Dyn* **231**, 214-20.
- Hartigan, N., Garrigue-Antar, L. and Kadler, K. E.** (2003). Bone morphogenetic protein-1 (BMP-1). Identification of the minimal domain structure for procollagen C-proteinase activity. *J Biol Chem* **278**, 18045-9.
- He, X., Saint-Jeannet, J. P., Woodgett, J. R., Varmus, H. E. and Dawid, I. B.** (1995). Glycogen synthase kinase-3 and dorsoventral patterning in *Xenopus* embryos. *Nature* **374**, 617-22.
- Heasman, J., Crawford, A., Goldstone, K., Garner-Hamrick, P., Gumbiner, B., McCrea, P., Kintner, C., Noro, C. Y. and Wylie, C.** (1994). Overexpression of cadherins and underexpression of beta-catenin inhibit dorsal mesoderm induction in early *Xenopus* embryos. *Cell* **79**, 791-803.
- Hemmati-Brivanlou, A., Kelly, O. G. and Melton, D. A.** (1994). Follistatin, an antagonist of activin, is expressed in the Spemann organizer and displays direct neuralizing activity. *Cell* **77**, 283-95.
- Hemmati-Brivanlou, A. and Thomsen, G. H.** (1995). Ventral mesodermal patterning in *Xenopus* embryos: expression patterns and activities of BMP-2 and BMP-4. *Dev Genet* **17**, 78-89.
- Henchcliffe, C., Garcia-Alonso, L., Tang, J. and Goodman, C. S.** (1993). Genetic analysis of laminin A reveals diverse functions during morphogenesis in *Drosophila*. *Development* **118**, 325-37.
- Hill, C. S.** (2001). TGF-beta signalling pathways in early *Xenopus* development. *Curr Opin Genet Dev* **11**, 533-40.

- Hojima, Y., van der Rest, M. and Prockop, D. J. (1985). Type I procollagen carboxyl-terminal proteinase from chick embryo tendons. Purification and characterization. *J Biol Chem* **260**, 15996-6003.
- Holley, S. A., Jackson, P. D., Sasai, Y., Lu, B., De Robertis, E. M., Hoffmann, F. M. and Ferguson, E. L. (1995). A conserved system for dorsal-ventral patterning in insects and vertebrates involving sog and chordin. *Nature* **376**, 249-53.
- Horb, M. E. and Slack, J. M. (2001). Endoderm specification and differentiation in *Xenopus* embryos. *Dev Biol* **236**, 330-43.
- Hornstra, I. K., Birge, S., Starcher, B., Bailey, A. J., Mecham, R. P. and Shapiro, S. D. (2003). Lysyl oxidase is required for vascular and diaphragmatic development in mice. *J Biol Chem* **278**, 14387-93.
- Hulmes, D. J., Mould, A. P. and Kessler, E. (1997). The CUB domains of procollagen C-proteinase enhancer control collagen assembly solely by their effect on procollagen C-proteinase/bone morphogenetic protein-1. *Matrix Biol* **16**, 41-5.
- Hwang, S. P., Partin, J. S. and Lennarz, W. J. (1994). Characterization of a homolog of human bone morphogenetic protein 1 in the embryo of the sea urchin, *Strongylocentrotus purpuratus*. *Development* **120**, 559-68.
- Hyde, C. E. and Old, R. W. (2000). Regulation of the early expression of the *Xenopus* nodal-related 1 gene, *Xnr1*. *Development* **127**, 1221-1229.
- Imamura, Y., Steiglitz, B. M. and Greenspan, D. S. (1998). Bone morphogenetic protein-1 processes the NH₂-terminal propeptide, and a furin-like proprotein convertase processes the COOH-terminal propeptide of pro- α 1(V) collagen. *J Biol Chem* **273**, 27511-7.
- Irish, V. F. and Gelbart, W. M. (1987). The decapentaplegic gene is required for dorsal-ventral patterning of the *Drosophila* embryo. *Genes Dev* **1**, 868-79.
- Isaacs, H. V. (1997). New perspectives on the role of the fibroblast growth factor family in amphibian development. *Cell Mol Life Sci* **53**, 350-61.
- Isaacs, H. V., Pownall, M. E. and Slack, J. M. (1994). eFGF regulates *Xbra* expression during *Xenopus* gastrulation. *Embo J* **13**, 4469-81.
- Isaacs, H. V., Pownall, M. E. and Slack, J. M. (1995). eFGF is expressed in the dorsal midline of *Xenopus laevis*. *Int J Dev Biol* **39**, 575-9.
- Janitz, M., Heiser, V., Bottcher, U., Landt, O. and Lauster, R. (1998). Three alternatively spliced variants of the gene coding for the human bone morphogenetic protein-1. *J Mol Med* **76**, 141-6.
- Jeay, S., Pianetti, S., Kagan, H. M. and Sonenshein, G. E. (2003). Lysyl oxidase inhibits ras-mediated transformation by preventing activation of NF-kappa B. *Mol Cell Biol* **23**, 2251-63.
- Jones, C. M., Kuehn, M. R., Hogan, B. L., Smith, J. C. and Wright, C. V. (1995). Nodal-related signals induce axial mesoderm and dorsalize mesoderm during gastrulation. *Development* **121**, 3651-62.
- Jones, C. M., Lyons, K. M., Lapan, P. M., Wright, C. V. and Hogan, B. L. (1992). DVR-4 (bone morphogenetic protein-4) as a posterior-ventralizing factor in *Xenopus* mesoderm induction. *Development* **115**, 639-47.
- Jones, C. M. and Smith, J. C. (1998). Establishment of a BMP-4 morphogen gradient by long-range inhibition. *Dev Biol* **194**, 12-7.
- Joseph, E. M. and Melton, D. A. (1997). *Xnr4*: a *Xenopus* nodal-related gene expressed in the Spemann organizer. *Dev Biol* **184**, 367-72.
- Jourdan-Le Saux, C., Tronecker, H., Bogic, L., Bryant-Greenwood, G. D., Boyd, C. D. and Csiszar, K. (1999). The LOXL2 gene encodes a new lysyl oxidase-like

protein and is expressed at high levels in reproductive tissues. *J Biol Chem* **274**, 12939-44.

Kagan, H. M. (1986). In *Biology of the Extracellular Matrix*, (ed. R. P. Mecham), pp. 321-398. New York: Academic Press.

Kao, K. R. and Elinson, R. P. (1989). Dorsalization of mesoderm induction by lithium. *Dev Biol* **132**, 81-90.

Keller, R., Davidson, L. A. and Shook, D. R. (2003). How we are shaped: the biomechanics of gastrulation. *Differentiation* **71**, 171-205.

Kenyon, K., Modi, W. S., Contente, S. and Friedman, R. M. (1993). A novel human cDNA with a predicted protein similar to lysyl oxidase maps to chromosome 15q24-q25. *J Biol Chem* **268**, 18435-7.

Kessler, E. and Adar, R. (1989). Type I procollagen C-proteinase from mouse fibroblasts. Purification and demonstration of a 55-kDa enhancer glycoprotein. *Eur J Biochem* **186**, 115-21.

Kessler, E., Fichard, A., Chanut-Delalande, H., Brusel, M. and Ruggiero, F. (2001). Bone morphogenetic protein-1 (BMP-1) mediates C-terminal processing of procollagen V homotrimer. *J Biol Chem* **276**, 27051-7.

Kessler, E., Takahara, K., Biniaminov, L., Brusel, M. and Greenspan, D. S. (1996). Bone morphogenetic protein-1: the type I procollagen C-proteinase. *Science* **271**, 360-2.

Khokha, M. K., Yeh, J., Grammer, T. C. and Harland, R. M. (2005). Depletion of three BMP antagonists from Spemann's organizer leads to a catastrophic loss of dorsal structures. *Dev Cell* **8**, 401-11.

Kim, M. S., Kim, S. S., Jung, S. T., Park, J. Y., Yoo, H. W., Ko, J., Csiszar, K., Choi, S. Y. and Kim, Y. (2003). Expression and purification of enzymatically active forms of the human lysyl oxidase-like protein 4. *J Biol Chem* **278**, 52071-4.

Kimelman, D. and Kirschner, M. (1987). Synergistic induction of mesoderm by FGF and TGF-beta and the identification of an mRNA coding for FGF in the early *Xenopus* embryo. *Cell* **51**, 869-77.

Kimelman, D. and Maas, A. (1992). Induction of dorsal and ventral mesoderm by ectopically expressed *Xenopus* basic fibroblast growth factor. *Development* **114**, 261-9.

Kirschmann, D. A., Seftor, E. A., Fong, S. F., Nieva, D. R., Sullivan, C. M., Edwards, E. M., Sommer, P., Csiszar, K. and Hendrix, M. J. (2002). A molecular role for lysyl oxidase in breast cancer invasion. *Cancer Res* **62**, 4478-83.

Klein, P. S. and Melton, D. A. (1996). A molecular mechanism for the effect of lithium on development. *Proc Natl Acad Sci U S A* **93**, 8455-9.

Koch, M., Laub, F., Zhou, P., Hahn, R. A., Tanaka, S., Burgeson, R. E., Gerecke, D. R., Ramirez, F. and Gordon, M. K. (2003). Collagen XXIV, a vertebrate fibrillar collagen with structural features of invertebrate collagens: selective expression in developing cornea and bone. *J Biol Chem* **278**, 43236-44.

Kofron, M., Demel, T., Xanthos, J., Lohr, J., Sun, B., Sive, H., Osada, S., Wright, C., Wylie, C. and Heasman, J. (1999). Mesoderm induction in *Xenopus* is a zygotic event regulated by maternal VegT via TGFbeta growth factors. *Development* **126**, 5759-70.

Larrain, J., Bachiller, D., Lu, B., Agius, E., Piccolo, S. and De Robertis, E. M. (2000). BMP-binding modules in chordin: a model for signalling regulation in the extracellular space. *Development* **127**, 821-30.

- Larrain, J., Oelgeschlager, M., Ketpura, N. I., Reversade, B., Zakin, L. and De Robertis, E. M.** (2001). Proteolytic cleavage of Chordin as a switch for the dual activities of Twisted gastrulation in BMP signaling. *Development* **128**, 4439-47.
- Lee, S. T., Kessler, E. and Greenspan, D. S.** (1990). Analysis of site-directed mutations in human pro-alpha 2(I) collagen which block cleavage by the C-proteinase. *J Biol Chem* **265**, 21992-6.
- Lepage, T., Ghiglione, C. and Gache, C.** (1992). Spatial and temporal expression pattern during sea urchin embryogenesis of a gene coding for a protease homologous to the human protein BMP-1 and to the product of the Drosophila dorsal-ventral patterning gene tolloid. *Development* **114**, 147-63.
- Li, S. W., Sieron, A. L., Fertala, A., Hojima, Y., Arnold, W. V. and Prockop, D. J.** (1996). The C-proteinase that processes procollagens to fibrillar collagens is identical to the protein previously identified as bone morphogenic protein-1. *Proc Natl Acad Sci U S A* **93**, 5127-30.
- Li, W., Nellaiappan, K., Strassmaier, T., Graham, L., Thomas, K. M. and Kagan, H. M.** (1997). Localization and activity of lysyl oxidase within nuclei of fibrogenic cells. *Proc Natl Acad Sci U S A* **94**, 12817-22.
- Lin, J. J., Maeda, R., Ong, R. C., Kim, J., Lee, L. M., Kung, H. and Maeno, M.** (1997). XBMP-1B (Xtld), a Xenopus homolog of dorso-ventral polarity gene in Drosophila, modifies tissue phenotypes of ventral explants. *Dev Growth Differ* **39**, 43-51.
- Liu, Q. R., Hattar, S., Endo, S., MacPhee, K., Zhang, H., Cleary, L. J., Byrne, J. H. and Eskin, A.** (1997). A developmental gene (Tolloid/BMP-1) is regulated in Aplysia neurons by treatments that induce long-term sensitization. *J Neurosci* **17**, 755-64.
- Liu, X., Zhao, Y., Gao, J., Pawlyk, B., Starcher, B., Spencer, J. A., Yanagisawa, H., Zuo, J. and Li, T.** (2004). Elastic fiber homeostasis requires lysyl oxidase-like 1 protein. *Nat Genet* **36**, 178-82.
- Loechel, F., Overgaard, M. T., Oxvig, C., Albrechtsen, R. and Wewer, U. M.** (1999). Regulation of human ADAM 12 protease by the prodomain. Evidence for a functional cysteine switch. *J Biol Chem* **274**, 13427-33.
- Lorent, K., Overbergh, L., Delabie, J., Van Leuven, F. and Van den Berghe, H.** (1994). Distribution of mRNA coding for alpha-2-macroglobulin, the murinoglobulins, the alpha-2-macroglobulin receptor and the alpha-2-macroglobulin receptor associated protein during mouse embryogenesis and in adult tissues. *Differentiation* **55**, 213-23.
- Maeno, M., Xue, Y., Wood, T. I., Ong, R. C. and Kung, H. F.** (1993). Cloning and expression of cDNA encoding Xenopus laevis bone morphogenetic protein-1 during early embryonic development. *Gene* **134**, 257-61.
- Maki, J. M. and Kivirikko, K. I.** (2001). Cloning and characterization of a fourth human lysyl oxidase isoenzyme. *Biochem J* **355**, 381-7.
- Maki, J. M., Rasanen, J., Tikkanen, H., Sormunen, R., Makikallio, K., Kivirikko, K. I. and Soininen, R.** (2002). Inactivation of the lysyl oxidase gene Lox leads to aortic aneurysms, cardiovascular dysfunction, and perinatal death in mice. *Circulation* **106**, 2503-9.
- Maki, J. M., Tikkanen, H. and Kivirikko, K. I.** (2001). Cloning and characterization of a fifth human lysyl oxidase isoenzyme: the third member of the lysyl oxidase-related subfamily with four scavenger receptor cysteine-rich domains. *Matrix Biol* **20**, 493-6.

- Marques, G., Musacchio, M., Shimell, M. J., Wunnenberg-Stapleton, K., Cho, K. W. and O'Connor, M. B.** (1997). Production of a DPP activity gradient in the early *Drosophila* embryo through the opposing actions of the SOG and TLD proteins. *Cell* **91**, 417-26.
- Mason, E. D., Konrad, K. D., Webb, C. D. and Marsh, J. L.** (1994). Dorsal midline fate in *Drosophila* embryos requires twisted gastrulation, a gene encoding a secreted protein related to human connective tissue growth factor. *Genes Dev* **8**, 1489-501.
- Mayne, R., Brewton, R. G., Mayne, P. M. and Baker, J. R.** (1993). Isolation and characterization of the chains of type V/type XI collagen present in bovine vitreous. *J Biol Chem* **268**, 9381-6.
- Medeck, R. J., Sosa, S., Morris, N. and Oxford, J. T.** (2003). BMP-1-mediated proteolytic processing of alternatively spliced isoforms of collagen type XI. *Biochem J* **376**, 361-8.
- Menendez, M., Gasset, M., Laynez, J., Lopez-Zumel, C., Usobiaga, P., Topfer-Petersen, E. and Calvete, J. J.** (1995). Analysis of the structural organization and thermal stability of two spermadhesins. Calorimetric, circular dichroic and Fourier-transform infrared spectroscopic studies. *Eur J Biochem* **234**, 887-96.
- Mentzer, R. K., Smith, N. D., Poch, G. and Dawson, D. A.** (1999). Combined osteolathyrin effects of beta-aminopropionitrile and penicillamine on *Xenopus* embryos: statistical comparison with dose-addition and independence. *Drug Chem Toxicol* **22**, 359-74.
- Miller, J. R., Rowning, B. A., Larabell, C. A., Yang-Snyder, J. A., Bates, R. L. and Moon, R. T.** (1999). Establishment of the dorsal-ventral axis in *Xenopus* embryos coincides with the dorsal enrichment of dishevelled that is dependent on cortical rotation. *J Cell Biol* **146**, 427-37.
- Miner, J. H., Li, C., Mudd, J. L., Go, G. and Sutherland, A. E.** (2004). Compositional and structural requirements for laminin and basement membranes during mouse embryo implantation and gastrulation. *Development* **131**, 2247-56.
- Molnar, J., Ujfaludi, Z., Fong, S. F., Bollinger, J. A., Waro, G., Fogelgren, B., Dooley, D. M., Mink, M. and Csiszar, K.** (2005). *Drosophila* lysyl oxidases Dmlox1-1 and Dmlox1-2 are differentially expressed and the active DmLOXL-1 influences gene expression and development. *J Biol Chem* **280**, 22977-85.
- Moon, R. T. and Kimelman, D.** (1998). From cortical rotation to organizer gene expression: toward a molecular explanation of axis specification in *Xenopus*. *Bioessays* **20**, 536-45.
- Moreno, M., Munoz, R., Aroca, F., Labarca, M., Brandan, E. and Larrain, J.** (2005). Biglycan is a new extracellular component of the Chordin-BMP4 signaling pathway. *Embo J* **24**, 1397-405.
- Nakatsuji, N.** (1986). Presumptive mesoderm cells from *Xenopus laevis* gastrulae attach to and migrate on substrata coated with fibronectin or laminin. *J Cell Sci* **86**, 109-18.
- Narita, R., Yamashita, H., Goto, A., Imai, H., Ichihara, S., Mori, H. and Kitagawa, Y.** (2004). Syndecan-dependent binding of *Drosophila* hemocytes to laminin alpha3/5 chain LG4-5 modules: potential role in sessile hemocyte islets formation. *FEBS Lett* **576**, 127-32.
- Nellaiappan, K., Risitano, A., Liu, G., Nicklas, G. and Kagan, H. M.** (2000). Fully processed lysyl oxidase catalyst translocates from the extracellular space into nuclei of aortic smooth-muscle cells. *J Cell Biochem* **79**, 576-82.

- Nguyen, T., Jamal, J., Shimell, M. J., Arora, K. and O'Connor, M. B.** (1994). Characterization of tolloid-related-1: a BMP-1-like product that is required during larval and pupal stages of *Drosophila* development. *Dev Biol* **166**, 569-86.
- Nieuwkoop, P. D.** (1969). The formation of mesoderm in Urodelean amphibians. I. Induction by the endoderm. *Roux's Arch Devl Biol* **162**, 341-373.
- Noramly, S., Zimmerman, L., Cox, A., Aloise, R., Fisher, M. and Grainger, R. M.** (2005). A gynogenetic screen to isolate naturally occurring recessive mutations in *Xenopus tropicalis*. *Mech Dev* **122**, 273-87.
- Oelgeschlager, M., Kuroda, H., Reversade, B. and De Robertis, E. M.** (2003). Chordin is required for the Spemann organizer transplantation phenomenon in *Xenopus* embryos. *Dev Cell* **4**, 219-30.
- Oelgeschlager, M., Larrain, J., Geissert, D. and De Robertis, E. M.** (2000). The evolutionarily conserved BMP-binding protein Twisted gastrulation promotes BMP signalling. *Nature* **405**, 757-63.
- Ogata, Y., Pratta, M. A., Nagase, H. and Arner, E. C.** (1992). Matrix metalloproteinase 9 (92-kDa gelatinase/type IV collagenase) is induced in rabbit articular chondrocytes by cotreatment with interleukin 1 beta and a protein kinase C activator. *Exp Cell Res* **201**, 245-9.
- Okada, T. S.** (1957). The Pluripotency of the Pharyngeal Primordium in Urodelean Neurulae. *J Embryol Exp Morphol* **3**, 438-448.
- Okada, T. S.** (1960). Epithelio-mesenchymal relationships in the regional differentiation of the digestive tract in the amphibian embryo. *Roux' Arch. f. Entw. mech.* **152**, 1-21.
- Osada, S. I. and Wright, C. V.** (1999). *Xenopus* nodal-related signaling is essential for mesendodermal patterning during early embryogenesis. *Development* **126**, 3229-40.
- Ottani, V., Martini, D., Franchi, M., Ruggeri, A. and Raspanti, M.** (2002). Hierarchical structures in fibrillar collagens. *Micron* **33**, 587-96.
- Pace, J. M., Corrado, M., Missero, C. and Byers, P. H.** (2003). Identification, characterization and expression analysis of a new fibrillar collagen gene, COL27A1. *Matrix Biol* **22**, 3-14.
- Padgett, R. W., St Johnston, R. D. and Gelbart, W. M.** (1987). A transcript from a *Drosophila* pattern gene predicts a protein homologous to the transforming growth factor-beta family. *Nature* **325**, 81-4.
- Padgett, R. W., Wozney, J. M. and Gelbart, W. M.** (1993). Human BMP sequences can confer normal dorsal-ventral patterning in the *Drosophila* embryo. *Proc Natl Acad Sci U S A* **90**, 2905-9.
- Palamakumbura, A. H., Jeay, S., Guo, Y., Pischon, N., Sommer, P., Sonenshein, G. E. and Trackman, P. C.** (2004). The propeptide domain of lysyl oxidase induces phenotypic reversion of ras-transformed cells. *J Biol Chem* **279**, 40593-600.
- Panchenko, M. V., Stetler-Stevenson, W. G., Trubetskoy, O. V., Gacheru, S. N. and Kagan, H. M.** (1996). Metalloproteinase activity secreted by fibrogenic cells in the processing of prolysin oxidase. Potential role of procollagen C-proteinase. *J Biol Chem* **271**, 7113-9.
- Parsons, M. J., Pollard, S. M., Saude, L., Feldman, B., Coutinho, P., Hirst, E. M. and Stemple, D. L.** (2002). Zebrafish mutants identify an essential role for laminins in notochord formation. *Development* **129**, 3137-46.
- Pei, D.** (1999). CA-MMP: a matrix metalloproteinase with a novel cysteine array, but without the classic cysteine switch. *FEBS Lett* **457**, 262-70.

- Petropoulou, V., Garrigue-Antar, L. and Kadler, K. E.** (2005). Identification of the Minimal Domain Structure of Bone Morphogenetic Protein-1 (BMP-1) for Chordinase Activity: CHORDINASE ACTIVITY IS NOT ENHANCED BY PROCOLLAGEN C-PROTEINASE ENHANCER-1 (PCPE-1). *J Biol Chem* **280**, 22616-23.
- Pfleiderer, G., Zwiiling, R. and Sonneborn, H. H.** (1967). [On the evolution of endopeptidases, 3. A protease of molecular weight 11,000 and a trypsin-like fraction from *Astacus fluviatilis* Fabr]. *Hoppe Seylers Z Physiol Chem* **348**, 1319-31.
- Piccolo, S., Agius, E., Leyns, L., Bhattacharyya, S., Grunz, H., Bouwmeester, T. and De Robertis, E. M.** (1999). The head inducer Cerberus is a multifunctional antagonist of Nodal, BMP and Wnt signals. *Nature* **397**, 707-10.
- Piccolo, S., Agius, E., Lu, B., Goodman, S., Dale, L. and De Robertis, E. M.** (1997). Cleavage of Chordin by Xolloid metalloprotease suggests a role for proteolytic processing in the regulation of Spemann organizer activity. *Cell* **91**, 407-16.
- Piccolo, S., Sasai, Y., Lu, B. and De Robertis, E. M.** (1996). Dorsoventral patterning in *Xenopus*: inhibition of ventral signals by direct binding of chordin to BMP-4. *Cell* **86**, 589-98.
- Pickard, B. and Damjanovski, S.** (2004). Overexpression of the tissue inhibitor of metalloproteinase-3 during *Xenopus* embryogenesis affects head and axial tissue formation. *Cell Res* **14**, 389-99.
- Prockop, D. J. and Hulmes, D. J.** (1994). Assembly of Collagen Fibrils *de Novo* from Soluble Precursors: Polymerisation and Copolymerisation of Procollagen, pN-Collagen, and Mutated Collagens. In *Extracellular Matrix Assembly and Structure*, (ed. P. D. Yurchenco D. E. Birk and R. P. Mecham): Academic Press.
- Prockop, D. J. and Kivirikko, K. I.** (1995). Collagens: molecular biology, diseases, and potentials for therapy. *Annu Rev Biochem* **64**, 403-34.
- Rattenholl, A., Pappano, W. N., Koch, M., Keene, D. R., Kadler, K. E., Sasaki, T., Timpl, R., Burgeson, R. E., Greenspan, D. S. and Bruckner-Tuderman, L.** (2002). Proteinases of the bone morphogenetic protein-1 family convert procollagen VII to mature anchoring fibril collagen. *J Biol Chem* **277**, 26372-8.
- Reynolds, S. D., Angerer, L. M., Palis, J., Nasir, A. and Angerer, R. C.** (1992). Early mRNAs, spatially restricted along the animal-vegetal axis of sea urchin embryos, include one encoding a protein related to tolloid and BMP-1. *Development* **114**, 769-86.
- Riggin, G. W. and Shultz, T. W.** (1986). Teratogenic effects of benzoyl hydrazine on frog embryos. *Trans Am Microsc Soc* **105**, 197-210.
- Ross, J. J., Shimmi, O., Vilmos, P., Petryk, A., Kim, H., Gaudenz, K., Hermanson, S., Ekker, S. C., O'Connor, M. B. and Marsh, J. L.** (2001). Twisted gastrulation is a conserved extracellular BMP antagonist. *Nature* **410**, 479-83.
- Saito, H., Papaconstantinou, J., Sato, H. and Goldstein, S.** (1997). Regulation of a novel gene encoding a lysyl oxidase-related protein in cellular adhesion and senescence. *J Biol Chem* **272**, 8157-60.
- Sampath, T. K., Rashka, K. E., Doctor, J. S., Tucker, R. F. and Hoffmann, F. M.** (1993). Drosophila transforming growth factor beta superfamily proteins induce endochondral bone formation in mammals. *Proc Natl Acad Sci U S A* **90**, 6004-8.
- Sandell, L. J., Morris, N., Robbins, J. R. and Goldring, M. B.** (1991). Alternatively spliced type II procollagen mRNAs define distinct populations of cells during vertebral development: differential expression of the amino-propeptide. *J Cell Biol* **114**, 1307-19.

- Sarrias, M. R., Gronlund, J., Padilla, O., Madsen, J., Holmskov, U. and Lozano, F.** (2004). The Scavenger Receptor Cysteine-Rich (SRCR) domain: an ancient and highly conserved protein module of the innate immune system. *Crit Rev Immunol* **24**, 1-37.
- Sasai, Y., Lu, B., Piccolo, S. and De Robertis, E. M.** (1996). Endoderm induction by the organizer-secreted factors chordin and noggin in *Xenopus* animal caps. *Embo J* **15**, 4547-55.
- Sasai, Y., Lu, B., Steinbeisser, H., Geissert, D., Gont, L. K. and De Robertis, E. M.** (1994). *Xenopus* chordin: a novel dorsalizing factor activated by organizer-specific homeobox genes. *Cell* **79**, 779-90.
- Schönherr, E., Witsch-Prehm, P., Harrach, B., Robenek, H., Rauterberg, J. and Kresse, H.** (1995). Interaction of biglycan with type I collagen. *J Biol Chem* **270**, 2776-83.
- Schulte-Merker, S., Lee, K. J., McMahon, A. P. and Hammerschmidt, M.** (1997). The zebrafish organizer requires chordin. *Nature* **387**, 862-3.
- Schulte-Merker, S. and Smith, J. C.** (1995). Mesoderm formation in response to Brachyury requires FGF signalling. *Curr Biol* **5**, 62-7.
- Schultz, T. W., Dumont, J. N. and Epler, R. G.** (1985). The embryotoxic and osteolathrogenic effects of semicarbazide. *Toxicology* **36**, 183-98.
- Scott, I. C., Blitz, I. L., Pappano, W. N., Imamura, Y., Clark, T. G., Steiglit, B. M., Thomas, C. L., Maas, S. A., Takahara, K., Cho, K. W. et al.** (1999). Mammalian BMP-1/Tolloid-related metalloproteinases, including novel family member mammalian Tolloid-like 2, have differential enzymatic activities and distributions of expression relevant to patterning and skeletogenesis. *Dev Biol* **213**, 283-300.
- Scott, I. C., Blitz, I. L., Pappano, W. N., Maas, S. A., Cho, K. W. and Greenspan, D. S.** (2001). Homologues of Twisted gastrulation are extracellular cofactors in antagonism of BMP signalling. *Nature* **410**, 475-8.
- Scott, I. C., Imamura, Y., Pappano, W. N., Troedel, J. M., Recklies, A. D., Roughley, P. J. and Greenspan, D. S.** (2000). Bone morphogenetic protein-1 processes biglycan. *J Biol Chem* **275**, 30504-11.
- Seidah, N. G. and Chretien, M.** (1999). Proprotein and prohormone convertases: a family of subtilases generating diverse bioactive polypeptides. *Brain Res* **848**, 45-62.
- Serpe, M., Ralston, A., Blair, S. S. and O'Connor, M. B.** (2005). Matching catalytic activity to developmental function: tolloid-related processes Sog in order to help specify the posterior crossvein in the *Drosophila* wing. *Development* **132**, 2645-56.
- Seufert, D. W., Hanken, J. and Klymkowsky, M. W.** (1994). Type II collagen distribution during cranial development in *Xenopus laevis*. *Anat Embryol (Berl)* **189**, 81-9.
- Shimell, M. J., Ferguson, E. L., Childs, S. R. and O'Connor, M. B.** (1991). The *Drosophila* dorsal-ventral patterning gene tolloid is related to human bone morphogenetic protein 1. *Cell* **67**, 469-81.
- Sieron, A. L., Tretiakova, A., Jameson, B. A., Segall, M. L., Lund-Katz, S., Khan, M. T., Li, S. and Stocker, W.** (2000). Structure and function of procollagen C-proteinase (mTolloid) domains determined by protease digestion, circular dichroism, binding to procollagen type I, and computer modeling. *Biochemistry* **39**, 3231-9.
- Sigrist, C. J., Cerutti, L., Hulo, N., Gattiker, A., Falquet, L., Pagni, M., Bairoch, A. and Bucher, P.** (2002). PROSITE: a documented database using patterns and profiles as motif descriptors. *Brief Bioinform* **3**, 265-74.

- Slack, J. M., Darlington, B. G., Heath, J. K. and Godsave, S. F.** (1987). Mesoderm induction in early *Xenopus* embryos by heparin-binding growth factors. *Nature* **326**, 197-200.
- Smith, J. C., Price, B. M., Van Nimmen, K. and Huylebroeck, D.** (1990). Identification of a potent *Xenopus* mesoderm-inducing factor as a homologue of activin A. *Nature* **345**, 729-31.
- Smith, J. C. and Watt, F. M.** (1985). Biochemical specificity of *Xenopus* notochord. *Differentiation* **29**, 109-15.
- Smith, K. M., Gaultier, A., Cousin, H., Alfandari, D., White, J. M. and DeSimone, D. W.** (2002). The cysteine-rich domain regulates ADAM protease function in vivo. *J Cell Biol* **159**, 893-902.
- Smith, W. C. and Harland, R. M.** (1992). Expression cloning of noggin, a new dorsalizing factor localized to the Spemann organizer in *Xenopus* embryos. *Cell* **70**, 829-40.
- Smith, W. C., McKendry, R., Ribisi, S., Jr. and Harland, R. M.** (1995). A nodal-related gene defines a physical and functional domain within the Spemann organizer. *Cell* **82**, 37-46.
- Snawder, J. E. and Chambers, J. E.** (1990). Critical time periods and the effect of tryptophan in malathion-induced developmental defects in *Xenopus* embryos. *Life Sci* **46**, 1635-42.
- Snawder, J. E. and Chambers, J. E.** (1993). Osteolathrogenic effects of malathion in *Xenopus* embryos. *Toxicol Appl Pharmacol* **121**, 210-6.
- Sottrup-Jensen, L.** (1989). Alpha-macroglobulins: structure, shape, and mechanism of proteinase complex formation. *J Biol Chem* **264**, 11539-42.
- Sottrup-Jensen, L., Sand, O., Kristensen, L. and Fey, G. H.** (1989). The alpha-macroglobulin bait region. Sequence diversity and localization of cleavage sites for proteinases in five mammalian alpha-macroglobulins. *J Biol Chem* **264**, 15781-9.
- Spemann, H. and Mangold, H.** (1924). Über Induktion von Embryonalanlagen durch Implantation artfremder Organisatoren. *Roux' Arch. f. Entw. mech.* **100**, 599-638.
- Springman, E. B., Angleton, E. L., Birkedal-Hansen, H. and Van Wart, H. E.** (1990). Multiple modes of activation of latent human fibroblast collagenase: evidence for the role of a Cys73 active-site zinc complex in latency and a "cysteine switch" mechanism for activation. *Proc Natl Acad Sci U S A* **87**, 364-8.
- Srinivasan, S., Rashka, K. E. and Bier, E.** (2002). Creation of a Sog morphogen gradient in the *Drosophila* embryo. *Dev Cell* **2**, 91-101.
- Stambolic, V., Ruel, L. and Woodgett, J. R.** (1996). Lithium inhibits glycogen synthase kinase-3 activity and mimics wingless signalling in intact cells. *Curr Biol* **6**, 1664-8.
- Stemple, D. L.** (2005). Structure and function of the notochord: an essential organ for chordate development. *Development* **132**, 2503-12.
- Stenflo, J., Stenberg, Y. and Muranyi, A.** (2000). Calcium-binding EGF-like modules in coagulation proteinases: function of the calcium ion in module interactions. *Biochim Biophys Acta* **1477**, 51-63.
- Stocker, W., Breit, S., Sottrup-Jensen, L. and Zwillig, R.** (1991). alpha2-Macroglobulin from hemolymph of the freshwater crayfish *Astacus astacus*. *Comp Biochem Physiol [B]* **98**, 501-9.
- Stocker, W., Gomis-Ruth, F. X., Bode, W. and Zwillig, R.** (1993). Implications of the three-dimensional structure of astacin for the structure and function of the astacin family of zinc-endopeptidases. *Eur J Biochem* **214**, 215-31.

- Su, M. W., Suzuki, H. R., Bieker, J. J., Solursh, M. and Ramirez, F. (1991).** Expression of two nonallelic type II procollagen genes during *Xenopus laevis* embryogenesis is characterized by stage-specific production of alternatively spliced transcripts. *J Cell Biol* **115**, 565-75.
- Sun, B. I., Bush, S. M., Collins-Racie, L. A., LaVallie, E. R., DiBlasio-Smith, E. A., Wolfman, N. M., McCoy, J. M. and Sive, H. L. (1999).** derriere: a TGF-beta family member required for posterior development in *Xenopus*. *Development* **126**, 1467-82.
- Suzuki, A. (1995).** Bone morphogenetic protein acts as a ventral mesoderm modifier in early *xenopus* embryos. *Dev. Growth Differ.* **37**, 581-588.
- Suzuki, N., Nakatsuka, H., Mochizuki, M., Nishi, N., Kadoya, Y., Utani, A., Oishi, S., Fujii, N., Kleinman, H. K. and Nomizu, M. (2003).** Biological activities of homologous loop regions in the laminin alpha chain G domains. *J Biol Chem* **278**, 45697-705.
- Takahara, K., Kessler, E., Biniaminov, L., Brusel, M., Eddy, R. L., Jani-Sait, S., Shows, T. B. and Greenspan, D. S. (1994a).** Type I procollagen COOH-terminal proteinase enhancer protein: identification, primary structure, and chromosomal localization of the cognate human gene (PCOLCE). *J Biol Chem* **269**, 26280-5.
- Takahara, K., Lyons, G. E. and Greenspan, D. S. (1994b).** Bone morphogenetic protein-1 and a mammalian tolloid homologue (mTld) are encoded by alternatively spliced transcripts which are differentially expressed in some tissues. *J Biol Chem* **269**, 32572-8.
- Takahashi, S., Yokota, C., Takano, K., Tanegashima, K., Onuma, Y., Goto, J. and Asashima, M. (2000).** Two novel nodal-related genes initiate early inductive events in *Xenopus* Nieuwkoop center. *Development* **127**, 5319-29.
- Takayama, Y., Takada, F., Nowatari, M., Kawakami, M. and Matsu-ura, N. (1999).** Gene structure of the P100 serine-protease component of the human Rα-reactive factor. *Mol Immunol* **36**, 505-14.
- Talts, J. F., Sasaki, T., Miosge, N., Gohring, W., Mann, K., Mayne, R. and Timpl, R. (2000).** Structural and functional analysis of the recombinant G domain of the laminin alpha4 chain and its proteolytic processing in tissues. *J Biol Chem* **275**, 35192-9.
- Tang, C. and Klinman, J. P. (2001).** The catalytic function of bovine lysyl oxidase in the absence of copper. *J Biol Chem* **276**, 30575-8.
- Tang, S. S., Trackman, P. C. and Kagan, H. M. (1983).** Reaction of aortic lysyl oxidase with beta-aminopropionitrile. *J Biol Chem* **258**, 4331-8.
- Thiel, S., Vorup-Jensen, T., Stover, C. M., Schwaeble, W., Laursen, S. B., Poulsen, K., Willis, A. C., Eggleton, P., Hansen, S., Holmskov, U. et al. (1997).** A second serine protease associated with mannan-binding lectin that activates complement. *Nature* **386**, 506-10.
- Thomsen, G. H. and Melton, D. A. (1993).** Processed Vg1 protein is an axial mesoderm inducer in *Xenopus*. *Cell* **74**, 433-41.
- Tortorella, M. D., Arner, E. C., Hills, R., Easton, A., Korte-Sarfaty, J., Fok, K., Wittwer, A. J., Liu, R. Q. and Malfait, A. M. (2004).** Alpha2-macroglobulin is a novel substrate for ADAMTS-4 and ADAMTS-5 and represents an endogenous inhibitor of these enzymes. *J Biol Chem* **279**, 17554-61.
- Tsubota, Y., Mizushima, H., Hirosaki, T., Higashi, S., Yasumitsu, H. and Miyazaki, K. (2000).** Isolation and activity of proteolytic fragment of laminin-5 alpha3 chain. *Biochem Biophys Res Commun* **278**, 614-20.

- Tsubota, Y., Yasuda, C., Kariya, Y., Ogawa, T., Hirosaki, T., Mizushima, H. and Miyazaki, K.** (2005). Regulation of biological activity and matrix assembly of laminin-5 by COOH-terminal, LG4-5 domain of alpha3 chain. *J Biol Chem* **280**, 14370-7.
- Umans, L., Serneels, L., Overbergh, L., Lorent, K., Van Leuven, F. and Van den Berghe, H.** (1995). Targeted inactivation of the mouse alpha 2-macroglobulin gene. *J Biol Chem* **270**, 19778-85.
- Umans, L., Serneels, L., Overbergh, L., Stas, L. and Van Leuven, F.** (1999). alpha2-macroglobulin- and murinoglobulin-1- deficient mice. A mouse model for acute pancreatitis. *Am J Pathol* **155**, 983-93.
- Unsold, C., Pappano, W. N., Imamura, Y., Steiglitz, B. M. and Greenspan, D. S.** (2002). Biosynthetic processing of the pro-alpha 1(V)2pro-alpha 2(V) collagen heterotrimer by bone morphogenetic protein-1 and furin-like proprotein convertases. *J Biol Chem* **277**, 5596-602.
- Uzel, M. I., Scott, I. C., Babakhanlou-Chase, H., Palamakumbura, A. H., Pappano, W. N., Hong, H. H., Greenspan, D. S. and Trackman, P. C.** (2001). Multiple bone morphogenetic protein 1-related mammalian metalloproteinases process pro-lysyl oxidase at the correct physiological site and control lysyl oxidase activation in mouse embryo fibroblast cultures. *J Biol Chem* **276**, 22537-43.
- Vaughan, J. M. and Vale, W. W.** (1993). Alpha 2-macroglobulin is a binding protein of inhibin and activin. *Endocrinology* **132**, 2038-50.
- Veitch, D. P., Nokelainen, P., McGowan, K. A., Nguyen, T. T., Nguyen, N. E., Stephenson, R., Pappano, W. N., Keene, D. R., Spong, S. M., Greenspan, D. S. et al.** (2003). Mammalian tolloid metalloproteinase, and not matrix metalloprotease 2 or membrane type 1 metalloprotease, processes laminin-5 in keratinocytes and skin. *J Biol Chem* **278**, 15661-8.
- Visconti, R. P., Barth, J. L., Keeley, F. W. and Little, C. D.** (2003). Codistribution analysis of elastin and related fibrillar proteins in early vertebrate development. *Matrix Biol* **22**, 109-21.
- Wallis, R. and Dodd, R. B.** (2000). Interaction of mannose-binding protein with associated serine proteases: effects of naturally occurring mutations. *J Biol Chem* **275**, 30962-9.
- Wang, S. X., Mure, M., Medzihradszky, K. F., Burlingame, A. L., Brown, D. E., Dooley, D. M., Smith, A. J., Kagan, H. M. and Klinman, J. P.** (1996). A crosslinked cofactor in lysyl oxidase: redox function for amino acid side chains. *Science* **273**, 1078-84.
- Wardle, F. C., Welch, J. V. and Dale, L.** (1999). Bone morphogenetic protein 1 regulates dorsal-ventral patterning in early *Xenopus* embryos by degrading chordin, a BMP4 antagonist. *Mech Dev* **86**, 75-85.
- Weaver, C., Farr, G. H., 3rd, Pan, W., Rowning, B. A., Wang, J., Mao, J., Wu, D., Li, L., Larabell, C. A. and Kimelman, D.** (2003). GBP binds kinesin light chain and translocates during cortical rotation in *Xenopus* eggs. *Development* **130**, 5425-36.
- Wenstrup, R. J., Florer, J. B., Brunskill, E. W., Bell, S. M., Chervoneva, I. and Birk, D. E.** (2004). Type V collagen controls the initiation of collagen fibril assembly. *J Biol Chem* **279**, 53331-7.
- Wessel, G. M. and McClay, D. R.** (1987). Gastrulation in the sea urchin embryo requires the deposition of crosslinked collagen within the extracellular matrix. *Dev Biol* **121**, 149-65.

- Williams, P. H., Hagemann, A., Gonzalez-Gaitan, M. and Smith, J. C.** (2004). Visualizing long-range movement of the morphogen Xnr2 in the *Xenopus* embryo. *Curr Biol* **14**, 1916-23.
- Wisniewski, H. G., Burgess, W. H., Oppenheim, J. D. and Vilcek, J.** (1994). TSG-6, an arthritis-associated hyaluronan binding protein, forms a stable complex with the serum protein inter-alpha-inhibitor. *Biochemistry* **33**, 7423-9.
- Wolfman, N. M., McPherron, A. C., Pappano, W. N., Davies, M. V., Song, K., Tomkinson, K. N., Wright, J. F., Zhao, L., Sebald, S. M., Greenspan, D. S. et al.** (2003). Activation of latent myostatin by the BMP-1/tolloid family of metalloproteinases. *Proc Natl Acad Sci U S A* **100**, 15842-6.
- Wozney, J. M., Rosen, V., Celeste, A. J., Mitsock, L. M., Whitters, M. J., Kriz, R. W., Hewick, R. M. and Wang, E. A.** (1988). Novel regulators of bone formation: molecular clones and activities. *Science* **242**, 1528-34.
- Wu, J. J. and Eyre, D. R.** (1995). Structural analysis of cross-linking domains in cartilage type XI collagen. Insights on polymeric assembly. *J Biol Chem* **270**, 18865-70.
- Yamada, K., Takabatake, T. and Takeshima, K.** (2000). Isolation and characterization of three novel serine protease genes from *Xenopus laevis*. *Gene* **252**, 209-16.
- Yamashita, H., Goto, A., Kadowaki, T. and Kitagawa, Y.** (2004). Mammalian and *Drosophila* cells adhere to the laminin alpha4 LG4 domain through syndecans, but not glypicans. *Biochem J* **382**, 933-43.
- Yang, M. and Kurkinen, M.** (1998). Cloning and developmental regulation of tissue inhibitor of metalloproteinases-3 (TIMP3) in *Xenopus laevis* early embryos. *Gene* **211**, 95-100.
- Yang, M., Murray, M. T. and Kurkinen, M.** (1997). A novel matrix metalloproteinase gene (XMMP) encoding vitronectin-like motifs is transiently expressed in *Xenopus laevis* early embryo development. *J Biol Chem* **272**, 13527-33.
- Yasumasu, S., Mao, K. M., Sultana, F., Sakaguchi, H. and Yoshizaki, N.** (2005). Cloning of a quail homologue of hatching enzyme: its conserved function and additional function in egg envelope digestion. *Dev Genes Evol*, 1-10.
- Yiallourous, I., Grosse Berkhoff, E. and Stocker, W.** (2000). The roles of Glu93 and Tyr149 in astacin-like zinc peptidases. *FEBS Lett* **484**, 224-8.
- Yu, K., Srinivasan, S., Shimmi, O., Biehs, B., Rashka, K. E., Kimelman, D., O'Connor, M. B. and Bier, E.** (2000). Processing of the *Drosophila* Sog protein creates a novel BMP inhibitory activity. *Development* **127**, 2143-54.
- Zhang, C., Basta, T., Hernandez-Lagunas, L., Simpson, P., Stemple, D. L., Artinger, K. B. and Klymkowsky, M. W.** (2004). Repression of nodal expression by maternal B1-type SOXs regulates germ layer formation in *Xenopus* and zebrafish. *Dev Biol* **273**, 23-37.
- Zhang, J., Houston, D. W., King, M. L., Payne, C., Wylie, C. and Heasman, J.** (1998). The role of maternal VegT in establishing the primary germ layers in *Xenopus* embryos. *Cell* **94**, 515-24.
- Zheng, X., Chung, D., Takayama, T. K., Majerus, E. M., Sadler, J. E. and Fujikawa, K.** (2001). Structure of von Willebrand factor-cleaving protease (ADAMTS13), a metalloprotease involved in thrombotic thrombocytopenic purpura. *J Biol Chem* **276**, 41059-63.
- Zhu, Y., Oganessian, A., Keene, D. R. and Sandell, L. J.** (1999). Type IIA procollagen containing the cysteine-rich amino propeptide is deposited in the

extracellular matrix of prechondrogenic tissue and binds to TGF-beta1 and BMP-2. *J Cell Biol* **144**, 1069-80.

Zimmerman, L. B., De Jesus-Escobar, J. M. and Harland, R. M. (1996). The Spemann organizer signal noggin binds and inactivates bone morphogenetic protein 4. *Cell* **86**, 599-606.

Zusman, S. B., Sweeton, D. and Wieschaus, E. F. (1988). short gastrulation, a mutation causing delays in stage-specific cell shape changes during gastrulation in *Drosophila melanogaster*. *Dev Biol* **129**, 417-27.

Zwilling, R. (1994). Astacin and the Astacins: A Re-View. In *The Astacins, Structure and Function of a New Protein Family*, (ed. R. Zwilling and W. Stocker), pp. pp. 11 to 25. Hamburg: Verlag Dr. Kovac.

Appendices

This Appendix includes the nucleotide and, where applicable, predicted protein sequences of all ESTs described in this thesis. Also included are the sequences of all primers used to make the Xld CUB deletion constructs.

A. *Xlox*

B. *Xlox1-1*

C. *Xlox1-3*

D. *ProCollagens* 1. *proCol3 α 1*

2. *proCol5 α 1*

3. *proCol5 α 2*

4. *proCol11 α 1*

E. *XBgn-EST*

F. *Laminin α -like*

G. Criteria for scoring ventralised phenotypes adapted from the DAI scale

H. Primers used for Xld CUB deletions

Appendix A - Xlox Sequence

```

      10      20      30      40      50      60      70
GTCCGGAGGGACAAAGATGGGCTGCGCCCCGGGGAGAGTCTTTTTATTGCAGCTCGTCCTGTATTTACATT
      M G C A P G R V F L L Q L V L Y L H>

      80      90     100     110     120     130     140
GTC TTTTACTTGGGGACTGT CAGCAGCAAAGGAGCCCTGTGGGGCCGAGCTCCATGTGGAGACAGAGACT
C L L L G D C Q Q Q R S P V G P S S M W R Q R L>

      150     160     170     180     190     200     210
TCAGTGGCAGAGTAACGGGAGGGGTGTACAGTCTGCTGAGCCACGGCTCCGAGTACCAGCCGGCCAGGCGC
Q W Q S N G R V Y S L L S H G S E Y Q P A R R>

      220     230     240     250     260     270     280
AGGGAGCAGGAGGGCAGCGCGCCCCACTCTCCGCTTCTTCTGCTCAGCAGCCGCCACAACAACAGCGTTT
R E Q E G S A P H S P L L L L S S R H N N S V>

      290     300     310     320     330     340     350
CCCAGGGCACCGCGCACACAGAGCCCAACCTCCAGATCCGTAACCCCAAGTGCGGAGGGGCAGCAGCA
S Q G T G A H R A Q P P R S V T P S A E G Q Q Q>

      360     370     380     390     400     410     420
ACAACCTGCGAGCACTAACCCGGCAGCACGGGAAGACGCCGCCACCTCTCGCTTGTTGGTTCCAGGCG
Q P A S T N P A A R E D A A A T S R L W F Q A>

      430     440     450     460     470     480     490
ACAAGGCAAAAGAAACGCCGAGATGCCCGGGAGACAGAGGAGCGCAGGTGCCGCGCAGGAGAGGGCGTC
T R Q R N A G D A A G R Q R S A G A A Q E R R>

      500     510     520     530     540     550     560
AGGGCGCACAGGCGTCTCACACCCGAGTCAACGGGACGGACAGAGTGGAGGGAATGCAGGGGGACGACCC
Q G A Q A S H T R V N G T D R V E G M Q G D D P>

      570     580     590     600     610     620     630
GTATAACCCCTACAAGTACTCCGAGGATAATCCCTATTACAACCTACTATGACACCTATGAGAGATCCAGG
Y N P Y K Y S E D N P Y Y N Y Y D T Y E R S R>

      640     650     660     670     680     690     700
ACGAATGAGAGGCAGAGCGCGGCTATGGCACCAGGTACTTTCAGAACGGTCTTCCAGACTTGTTCCAG
T N E R Q R P G Y G T R Y F Q N G L P D L V P>

      710     720     730     740     750     760     770
ATCCTTACTACATT CAGGCGTCGATTTATGTT CAGAAAATGTCCATGTATAACCTGCGGTGTGCTGCTGA
D P Y Y I Q A S I Y V Q K M S M Y N L R C A A E>

      780     790     800     810     820     830     840
AGAAACTGCTTGGCAAGCTCAGCATACAGCTCAGATGTCAGAGACTATGATCAACGAGTGCTTCTCAGG
E N C L A S S A Y S S D V R D Y D Q R V L L R>

      850     860     870     880     890     900     910
TTCCACAAAAGAGTCAAAAACCAAGGTACGGCTGACTTCTTGCCAAGCAGACCACGATACACTTGGGAAT
F P Q R V K N Q G T A D F L P S R P R Y T W E>

      920     930     940     950     960     970     980
GGCACAGCTGT CACCAACATTATCATAGCATGGACGAGTTCAGCCATTATGATTGCTGGATGCCAGCAG
W H S C H Q H Y H S M D E F S H Y D L L D A S S>

      990     1000     1010     1020     1030     1040     1050
TCACAGAAGAGTTGCTGAGGGCCACAAAGCCAGCTTCTGCCTAGAAGATACCTCCTGTGATTATGGCTAT
H R R V A E G H K A S F C L E D T S C D Y G Y>

      1060     1070     1080     1090     1100     1110     1120
TACAGACGGTTTGCATGCACAGCTCATACACAGGGATTGAGTCCTGGTTGTTATGATACATATAATGCAG
Y R R F A C T A H T Q G L S P G C Y D T Y N A>

      1130     1140     1150     1160     1170     1180     1190
ATATTGATTGCCAGTGGATCGATATCAGATGTAAAACCTGGCAACTACATCTTAAAGGTCAGTGTGAA
D I D C Q W I D I T D V K P G N Y I L K V S V N>

      1200     1210     1220     1230     1240     1250     1260
TCCCAGCTACCAGGTACCAGAATCTGATTACTCCAACAATGTTGTGCGTTGTGATTTACGATACACTGGA
```

P S Y Q V P E S D Y S N N V V R C D L R Y T G>
 1270 1280 1290 1300 1310 1320 1330
 CATTACGTGTACACTTCCGGCTGCACAATTTCTCCTTAITAAATGCAAGGCAAGGGTTTATACTCAAGTG
 H Y V Y T S G C T I S P Y *>
 1340 1350 1360 1370 1380 1390 1400
 AAAGATGATTGGGATGTAACACATTTTAGGTGAAAGCATATGGACTTCCATGAAATGCAACTGGATTTG
 1410 1420 1430 1440 1450 1460 1470
 AATCACGTTTCAGAAAGTTAGGAGCATTGGGATGATCAGAAATTGTTGTACATTTCTCTTGTCTCCACGT
 1480 1490 1500 1510 1520 1530 1540
 ACAGACACTTCATGTACATGTAGAAGCGAGCTTTAAGAAGTCTGTTTATATACCAGCCAATAATAGTTAC
 1550 1560 1570 1580 1590 1600 1610
 ATTTATTTTTCAATTCTTAAAGAATGAATATTATATTGTATATCCTCTATAGAGTCTATAAGTACATA
 1620 1630 1640 1650 1660 1670 1680
 AAGCAATCTGACATGATCCATTGTGACGTTGAACATAGATGCCACTATGGAGAGAACACACCAGTGG
 1690 1700 1710 1720 1730 1740 1750
 ATAAAAGTGTAGTGTGAATAGTAGTTGGAGGACCTAATCAGAATGTTTGAATCCAAGTAAATTCATAGGA
 1760 1770 1780 1790 1800 1810 1820
 CTATGCCAAGGAATACCAACCTGAATAGTTATCATACAAAATGCTCCTCTTTTAAACAAAAAGGTGACA
 1830 1840 1850 1860 1870 1880 1890
 TTACAATGCAATACATGCCATGCTCTTCTGATTACATATTCTGTTGACCATCACAATTTTGGATCCACTT
 1900 1910 1920 1930 1940 1950 1960
 CCACTAGTGCCTTTAAAGCTGGACTACATCTCGTAGTATCCCCCAATTGTTGTATCAGTCATACACTGAG
 1970 1980 1990 2000 2010 2020 2030
 CCATTAATGACTGTGCAAGGATACCTCAAATGTTCTATCGGGCTTTTGGGTTGTCACCATATCAAAAGA
 2040 2050 2060 2070 2080 2090 2100
 CATTTCATAATTTCAGTTGCAGTAAATCTGCTCTGCCCTGAAGCATATACAGGTATGGGACATGTTATCCAGA
 2110 2120 2130 2140 2150 2160 2170
 ATGCTCAGGACCTGGGGTCTCCTGGATCTTTTGTAAATTTGGATCTTCATACCTTAAGTCTACTAGAAAA
 2180 2190 2200 2210 2220 2230 2240
 TCATGTAAACTTTAAATAAACCCAAATAGGCTGGTTTGTCTTCCAATAAGAATTAATTATATCTTAGTTGG
 2250 2260 2270 2280 2290
 GATCAACTACAAGGTACGGTTTATTATTACAGAGAAAAAAAAAAAAAAAA

Appendix B – Xloxl-1 Sequence

```

      10      20      30      40      50      60      70
GGAGTGACAGCAAGCGGGGTGCTGCTGAGCATGGCCCTTCCCCGGGCAACTCTGCCTATGGTTTGGGCTC
      M A F P G Q L C L W F W A>

      80      90     100     110     120     130     140
TTTCTGTGGCGCTGGGGCAGGAGGGTGACAGTTGGGGGCAGATGATCCAGTGGGAAGATAGTGGGAGGCG
L S V A L G Q E G D S W G Q M I Q W E D S G R R>

      150     160     170     180     190     200     210
CTATCGGCTACTAAATTCAAGCTCTGAGTACCAGCGCGCGGGGAGGTACCGGTGGATCCCGAGTCTTG
Y R L L N S G S E Y Q A A G G G T G G S R V L>

      220     230     240     250     260     270     280
CTGGATGGGATCCGTAGTGGTGGTTCGGACCTGCGCGCGCGGCAAGCACCTAGTTCCCCGCGTACTAGCT
L D G I R S G G S D L R R R Q A P S S P R T S>

      290     300     310     320     330     340     350
CCCAAACGGTCAGGGGGAATACCAGACACCCGTTTCGGCTTCGGACAGGTTCCCGACAACCTGGCGCAGCGT
S Q T V R G N T R H P F G F G Q V P D N W R S V>

      360     370     380     390     400     410     420
AAGCGAATCAGCAAGTACTGGGCGCTTCGTCCCTTCCGCAGGGGCTGGCGGTGGCGGCGGAAGGATACGC
S E S A S T G R F V P S A G A G G G G G R I R>

      430     440     450     460     470     480     490
CAGTCCTCTTCCAGGCTTCTGTGGGTACACAGTTTTCCTCAAGGCCCAACCCCTTTTGTTCCCCAA
Q S S S Q A S V G T Q F F P R P Q P P F V P Q>

      500     510     520     530     540     550     560
CTGACTTAAATCCTCAGGGATACGATGAGGGGTACGGGTACCAGAGAGCAGGCGGAGGGGAGGAAGCTA
T D L N P Q G Y D E G Y G Y Q R A G G G G G S Y>

      570     580     590     600     610     620     630
CTATGCGCAGTCATGGGCTGGTGGGTATGAAGACGCTTGGGAGGAGCCCTCCCCACCTTTTGACAAA
Y A Q S W A G G G Y E D A W E E P S P P F A Q>

      640     650     660     670     680     690     700
CCCCCTTATTTGGGATCCCCCTAACCTACCACCACAACAAGCTCCTCAAGCCAACCCAGTGGTGCCCC
P P Y F G I P P N L P P Q Q A P Q A N P V V P>

      710     720     730     740     750     760     770
AGGATGGACTGGATCGACGCTTTGCACATAGTCTGTTCGCGGTGAAGACCCCCAGTTGTAGCTAGTGA
Q D G L D R R F A H S L F R G E D P P V V A S E>

      780     790     800     810     820     830     840
GCCTGTCCGTCCTGGTTATGGTGGTGGGGGGGATACTGGGTTCCCCCTTTTGGAGGGGCGAGAGAGACA
P V R P G Y G G G G D T G F P P F G G A R E T>

      850     860     870     880     890     900     910
GATGGAGGATATTATGGGTCAACTAGACCAGAGCCAGTCCCTACCCCCAGCCAGAGTACCTAGTGCCCTTG
D G G Y Y G S T R P E P V L P P A R V P S A P>

      920     930     940     950     960     970     980
CCTCAGGAGGCCAAGCCAGGTAGAGCAACAAGGCAGAACCAGTGTGGGCAACGTATTCCGAGGCAACCA
A S G G Q A Q V E Q Q G R T S V G N V F R G N Q>

      990     1000    1010    1020    1030    1040    1050
AAATGGGAGAGGGCTTCTGATTGTGACCTGATCCTAGCTATGTTCAAGCTGCTACATACATTACGCGA
N G R G L P D L V P D P S Y V Q A A T Y I Q R>

      1060    1070    1080    1090    1100    1110    1120
GCTCATCTGTATTCTCTGAGGTGTGCGGCTGAAGAGAACTGTCTCTCCAGCTCTGCATACGCTGCTGAAG
A H L Y S L R C A A E E N C L S S A Y A A E>

      1130    1140    1150    1160    1170    1180    1190
CCACAGATTATGATGTTTCGAGTTCTTCTGCGATTTCACAAAGGGTGAAAAATCAGGGTACTGCTGACTT
A T D Y D V R V L L R F P Q R V K N Q G T A D F>

      1200    1210    1220    1230    1240    1250    1260
TCTGCCTACACGCCACGCCAGTCTGGGAATGGCACAGCTGCCACCAGCACTATCACAGTATGGATGAG
```

L P T R P R Q S W E W H S C H Q H Y H S M D E>
 1270 1280 1290 1300 1310 1320 1330
 TTCAGTCATTATGACTTGCTGGATGCTACCCACAGGCAGAAAAGTTGCAGAGGGACATAAAGCCAGCTTCT
 F S H Y D L L D A T T G R K V A E G H K A S F>
 1340 1350 1360 1370 1380 1390 1400
 GTTTAGAGGATACTACATGTGACTTTGGGAACTTAAAGAGATATGCATGCACTTCACACACTCAGGGTCT
 C L E D T T C D F G N L K R Y A C T S H T Q G L>
 1410 1420 1430 1440 1450 1460 1470
 GAGTCCAGGATGCTACGATACTTACAATGCAGATATAGACTGTCAGTGGATCGATATAACAGAGGTGAAA
 S P G C Y D T Y N A D I D C Q W I D I T E V K>
 1480 1490 1500 1510 1520 1530 1540
 CCCGAAACTATATACTAAAGGTTGTAGTGAATCCCAAGTATAAGGTTCTGGAGTCGGACTTCACCAATA
 P G N Y I L K V V V N P K Y K V L E S D F T N>
 1550 1560 1570 1580 1590 1600 1610
 ACGTGGTGCCTGTAAACATCCACTATACCGGACGTTATGCGTCTGCTACAAACTGCCGCATTACACAGTT
 N V V R C N I H Y T G R Y A S A T N C R I T Q F>
 1620 1630 1640 1650 1660 1670 1680
 TTGAGCAAGACTTTCAAGGACAAACGTACCTATGGGAGATTAAGTTCCTGCACCCCTGAACAAATCTGTTT
 *>
 1690 1700 1710 1720 1730 1740 1750
 ATTTTCCCTTTTACTCACACTGGGATTTGTATTAAATATTTGTGTAAAAGGATTTTATTTTTTCAAATAA
 1760 1770 1780 1790 1800 1810 1820
 AATCTATTTTCTAATTTTCATACAGATTCATGAATCAGAATAAACATCTAATGAGGGTAAAACAGCACAG
 1830 1840 1850 1860
 TCTTAAATAAAAACTTTTATCTCTTCAAAAAAAAAAAAAA

Appendix C - Xloxl-3 partial sequence

```

      10      20      30      40      50      60      70
TGGATCCCCGGGCTGCAGGAATTCGGCACGAGGGCTACTCGCTGCATGCGGTGACGGAAACGTGGTACT
W I P R A A G I R H E G Y S L H A V T E T W Y>

      80      90     100     110     120     130     140
GGGACGCCAGCAACACCACAGAGATGGTTTTGAGTGGTGTTCGCTGCACTGGAAGAGAAATGTCCTTGGGA
W D A S N T T E M V L S G V R C T G R E M S L E>

      150     160     170     180     190     200     210
ACAATGTAGCCACCATTTCGAGTGTCTGAGTGTCAAGAACTGGGACCCGGCACGCGGCTGGAGTGTATCTGC
Q C S H H S S V S C K N T G T R H A A G V I C>

      220     230     240     250     260     270     280
TCTGAGACGGCCTCAGACCTGGTGTCTGCACTCGTCTGGTACAGGAGACGGCCTACATCGAGGATCGGC
S E T A S D L V L H S S L V Q E T A Y I E D R>

      290     300     310     320     330     340     350
CCCTACACATGCTGTACTGTGCTGCTGAGGAGAACTGTCTCTCTAGTCCGCCCGCCATGCCAACTGGCC
P L H M L Y C A A E E N C L S S S A R H A N W P>

      360     370     380     390     400     410     420
TTATGGACACCGACGCTCTCCGCTTCTCATCCCAGATCCACAACATGGTCTGTCAGACTTCAGGCCT
Y G H R R L L R F S S Q I H N N G R A D F R P>

      430     440     450     460     470     480     490
AAGGCTGGTCGCCACTCTTGGGTCTGGCATGAGTGTACGCGGCATTATCACAGCATGGACATCTTCACCC
K A G R H S W V W H E C H G H Y H S M D I F T>

      500     510     520     530     540     550     560
ACTATGATATGCTAACGCCCAATGGTACCAAGGTGGCAGAAGGGCACAAAGCCAGCTTCTGTCTGGAAGA
H Y D M L T P N G T K V A E G H K A S F C L E D>

      570     580     590     600     610     620     630
CTCAGAGTGTTCAGGAATTAGTCAGTAAGAGATATGAATGTGCAAATTTTGGGGAGCAAGGAATCACAGTA
S E C Q E L V S K R Y E C A N F G E Q G I T V>

      640     650     660     670     680     690     700
GGATGCTGGGATTTATATCGACATGATATTGACTGCCAGTGGATTGATATAACAGATGTAAACCGGGAA
G C W D L Y R H D I D C Q W I D I T D V K P G>

      710     720     730     740     750     760     770
ACTACATCCTACAGGTTGTTATCAACCCAACTTTGAAGTGGCTGAGAGTGATTTCACCAATAACGCAAT
N Y I L Q V V I N P N F E V A E S D F T N N A M>

      780     790     800     810     820     830     840
GAAATGTAAGTAAATATGACGGACATCGAATCTGGATACACAACCTGCCATCTGGGTGACGCATTTCAGC
K C N C K Y D G H R I W I H N C H L G D A F S>

      850     860     870     880     890     900     910
GAAGAAGCAAATAAGCGTTTGTGAGCATACAGGGGAGAACTTAACAACCAGATACTTTGATGCCTGGTTTA
E E A N K R F E H Y R G E L N N Q I L *

      920     930     940     950     960     970     980
GTACCGGTCTGAAGACGCCCTCCTTTTTTGCAACAAGAAACAAGTTGGAGTAGAAACATGTGAACCTTTG

      990     1000     1010     1020     1030     1040     1050
TTGTTCTGGTAGTGAAGTATTGACTCCTGCTCACAGCACAGATTGGACTGCTTGTGTAGTGCCTTTTCC

      1060     1070     1080     1090     1100     1110     1120
CAGTGTGGGGCCACTATCGATGGTATGAATCGTTTTTGGGGGCGAGTGAAGGGAATTTTATATCCATTGTG

      1130     1140     1150     1160     1170     1180     1190
AAGAATCTGTCTCCGAGACATGAGTTTACCAAGTTGGGAAGGAACCTTGGAAAATCGCCCCCCTCTTT

      1200     1210     1220     1230     1240     1250     1260
CGTGTAGGGCAGTTTTGCTGAAGTAGATTGGAGGGTGCTTCCTTGCTTACCTGTTCTTCTGTTTTC
```

1270 1280 1290 1300 1310 1320 1330
CATTGAGGTCTGGTACGCTTGGTCCCCTGGTCAGAGTTGACTCAAAGGAGGGAAAACCTCTGTTGTGTGA

1340 1350 1360 1370 1380 1390 1400
GGACCAGTTTCTTCTATTAGAGAATCTAATAAGTGGATCTGAGTGCCCCAAGAAGCTTCTCTTACCTCCG

1410 1420 1430 1440 1450 1460 1470
GTATCCTCTGTAATGATTTGCTTTAAGTTGGGTCCAAACAATGAATCACAGGAATATTGAGATTTAATA

1480 1490 1500 1510 1520 1530 1540
GGCGGTTCTTTGACCCAGTATCTCCTGACCAGTGATCCTGCCTAGCCACTTCACATTGGCTGTAGACTTT

1550 1560 1570 1580 1590 1600 1610
GCTGATAAGTCGATCATTTACAAGGTGAAATCTGCTAGAAACAACAGAGCGGTGTTTAAATGTCTATATCA

1620 1630 1640 1650 1660
GCTGCGAATAAAGAAATGGTGGTTGTTTGTGAAAAAAAAAAAAAAAAAAAA

Appendix D1

Nucleotide sequence of proCola1III versus corresponding region of Hs proCola1III

HsCol3a1	2401	ATTGGTCTCTCTGGCCAGCTGGCCAGCCTGGAGATAAGGGTGAAGGTGG	2450
Col3a1 Sp6	1	AGGCTGG	7
		* * *	
HsCol3a1	2451	TGCC--CCCGGACTTCCAGGTATAGCTGGACCTCGTGGTAGCCCTGGTGA	2498
Col3a1 Sp6	8	TACCGGTCCGGAATCCCGGGATATCGTCGACCCACGC--GTCCGGGTGA	55
		* * * * *	
HsCol3a1	2499	GAGAGGTGAACTGGCCCTCCAGGACCTGCTGGTTTCCCTGGTGCTCCTG	2548
Col3a1 Sp6	56	ACGAGGAGAGTCTGGCCCTCTGGCCCTGCTGGCTTCCCTGGAGCCCTG	105
		* * * * *	
HsCol3a1	2549	GACAGAATGGTGAACCTGGTGGTAAAGGAGAAAGAGGGGCTCCGGGTGAG	2598
Col3a1 Sp6	106	GTCAGAATGGAGAACCTGGAGTCAAAGGTGAGAAAGGTGGTCTGGGGAA	155
		* * * * *	
HsCol3a1	2599	AAAGGTGAAGGAGGCCCTCTGGAGTTGCAGGACCCCTGGAGGTTCTGG	2648
Col3a1 Sp6	156	CGCGGAGAGTCTGGTAGTCTGGAGTTGGTGGCCCTCAAGGTGGTCTGG	205
		* * * * *	
HsCol3a1	2649	ACCTGCTGGTCTCTCTGGTCCCCAAGGTGTCAAAGGTGAACGTGGCAGTC	2698
Col3a1 Sp6	206	ACCTGCTGGTAATCCCGTCTCTCAAGGAGTAAAGGAGACCGTGGTTCTC	255
		* * * * *	
HsCol3a1	2699	CTGGTGGACCTGGTGTCTGCTGGCTTCCCTGGTGTCTGGTCTTCCTGGT	2748
Col3a1 Sp6	256	CAGCGGAGCTGGTGGCGCTGGTTTCCGGGTACTCGCGGTTGCCCTGGT	305
		* * * * *	
HsCol3a1	2749	CCTCCTGGTAGTAATGGTAACCCAGGACCCCGAGGTCCCAGCGGTTCTCC	2798
Col3a1 Sp6	306	CCTCCTGGATCTAATGGCAATGCAGGTCTCTGGCCCTCTGGAGGTCC	355
		* * * * *	
HsCol3a1	2799	AGGCAAGGATGGGCCCCAGGTCTGCGGGTAACACTGGTGTCTCTGGCA	2848
Col3a1 Sp6	356	TGGAAAAATGGACCTCCAGGTCCATCTGGCAATGTTGGTGTCCAGGAA	405
		* * * * *	
HsCol3a1	2849	GCCCTGGAGTGTCTGGACCAAAAGGTGATGCTGGCCAACAGGAGAGAAG	2898
Col3a1 Sp6	406	GTGCTGGTGCAGCAGGTCCAAGGGTGAACAGGCCAAGCAGGAGAGAGA	455
		* * * * *	
HsCol3a1	2899	GGATCGCTGGTGGCCAGGGCCACCAGGAGCTCCAGGCCACTTGGGAT	2948
Col3a1 Sp6	456	GGTGCACCTGGTTCTCGTGGTGAAGCTGGTGTCTCCAGGTCTCTCT	500
		* * * * *	
HsCol3a1	4751	TATAATAAATAAACTTCAACACTCTTTATGATAACAACACTGTGTTATAT	4800
Col3a1 T7	1	TTTTAGTTAAAAAAAACCACTGTGT	28
		* * * * *	
HsCol3a1	4801	TCTTTGAATCCTAGCCCATCTGCAGAGCAATGACTGTGCTCACCAGTAAA	4850
Col3a1 T7	29	GATTTTTT-TTTGGGCCAAGATGTATAGCACTGATCATTCAGGCA-TCAA	76
		* * * * *	
HsCol3a1	4851	AGATAACCTTTCTTTCTGA---AATAGTCAAATACGAAATTAGAAAAG--	4895
Col3a1 T7	77	ACTCAACCCATCTTTCTAACCAGATGACATTACAATACCAAAAAGAA	126
		* * * * *	
HsCol3a1	4896	CCCTCCCTATTTTAACTACCTCACTGGTCAGAAACACAGATTGTATTCT	4945
Col3a1 T7	127	CTTCCCTGTTTGAACCTACCTCA-----TCATGGACAGAATTGGG----	167
		* * * * *	
HsCol3a1	4946	ATGAGTCCCAGAGATGAAAAAATTTTATACGTTGATAAACTTATAAA	4995
Col3a1 T7	168	GGGAGTTTGATGAGGTAGAAAA-----CGGATTTCATTGGAATCAT---	209
		* * * * *	
HsCol3a1	4996	TTTCATTGATTAATCTCCTGGAAGATTGGTTTAAAAAGAAAAGTGAATG	5045
Col3a1 T7	210	TGCAATTAAATTAAG-CAAAAATGATGCAGTTGTAAGGGAATATGTGAGT	258
		* * * * *	
HsCol3a1	5046	CAAGAATTTAAAGAAATATTTTAAAGCCACAATATTTTAAATATTGGAT	5095
Col3a1 T7	259	CCCAAATT---ATTATATGATTATGGATATCAT-ATCTGAAGCGGTCT	303
		* * * * *	
HsCol3a1	5096	ATCAACTGCTTGTAAGGTGCTCCTCTTTTCTTGTCATTGCTGGTCAA	5145
Col3a1 T7	304	AGCAGCTT-----GGTGTCTTTATATTTTCTGGCATGTCTGGCAAA	345
		* * * * *	
HsCol3a1	5146	GATTACTAATATTTGGGAAGGCTTTAAAGAC-GCATGTTATGG-TGCTAA	5193
Col3a1 T7	346	GACCAGTTAAAAAGTAAATGATTTCAAAGATTATCCATCATGGGTGCTAA	395
		* * * * *	
HsCol3a1	5194	TGTACTTTTCACTTTTAACTCTAGATCAGAATTGTTGACTTGCATTGAGA	5243
Col3a1 T7	396	T-TGCTAATATGTTTCCATTCTAAAACAGCAACACAAAGAAATATTGAGT	444
		* * * * *	
HsCol3a1	5244	ACATAAATGCACAAAATCTGTACATGCTCCCATCAGAAAGATTCAATTGG	5293
Col3a1 T7	445	G-GTATGAACATGATGAAATAAAATGTTTCACAAAAAAGAAAAA	490
		* * * * *	

Appendix D2

Alignment of proCol5a1 versus the corresponding region of Rat proCol5a1

```
Rat Col 5a1 1851 AGTGATGGTGCCAGAGGCATGCCTGGACAAACAGGCCCAAGGGTGACCG 1900
proCol5a1 T7 1 GGTGACCG 8
*****

Rat Col 5a1 1901 TGGCTTTGATGGTCTGGCTGGGTTGCCGGGAGAGAAAGGCCATAGAGGTG 1950
proCol5a1 T7 9 GGGATTTCGATGGGTTGGCTGGTTTACCAGGCGAAAAGGGACATCGTGGGG 58
** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** **

Rat Col 5a1 1951 ACCCTGGTCCTTCTGGCCCGCCGGGACTCCCAGGAGATGATGGAGAAAGG 2000
proCol5a1 T7 59 AGTCTGGCCCAACACGGTCCCCCTGGGCCACCTGGAGAGGACGGAGAAAGG 108
* ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** * ** ** ** * ** ** ** * ** ** ** *

Rat Col 5a1 2001 GGTGACGATGGAGAAGTTGGGCCAGGGGGCTGCCTGGGGAGCCTGGACC 2050
proCol5a1 T7 109 GGTGATGATGGAGATGTTGGACCTAGAGGGTTGCCAGGTGAACCCGGACC 158
*****
***** ** ** ** ** ** ** ** ** ** ** ** ** * ** ** ** * ** ** ** * ** ** ** *

Rat Col 5a1 2051 ACGTGGTCTGCTTGGGGCCAAAAGGCCCCCTGGCCACCTGGACCTCCTG 2100
proCol5a1 T7 159 ACGAGGGTTGCTTGGACCTAAAGGACCTCAAGGGCCACCAGGATCTACTG 208
*** ** * ** ** ** * ** ** * ** * ** ** ** * ** ** ** *

Rat Col 5a1 2101 GTGTGACGGGTATGGATGGCCAGCCTGGCCCCAAAGGAAATGTGGGTCCC 2150
proCol5a1 T7 209 GTGTTCTTGGAAATGGACGGTCAAGTTGGCCCCAAAGGTAGTGTGGGGCCT 258
***** * ** ** ** * ** * ** ** ** * ** ** ** * ** ** ** *

Rat Col 5a1 2151 CAGGGAGAGCCTGGGCCCCAGGACAGCAGGGTAATCCTGGTGCCAGGG 2200
proCol5a1 T7 259 CAAGGTGAGCCAGGACCACCCGGGCAGCAAGGGAACCCAGGAGCACAGGG 308
** ** ** ** * ** ** * ** ** * ** ** * ** ** * ** ** * ** ** *

Rat Col 5a1 2201 TCTTCCCGGTCCCCAGGGTGCCATTGGTCCTCCAGGAGAAAAGGGTCCTT 2250
proCol5a1 T7 309 TCTTCCAGGTCTCTCAGGGTGCAATCGGCCACCAGGAGAAAAGGGCCCTT 358
***** ** ** ** * ** ** * ** ** * ** ** * ** ** * ** ** *

Rat Col 5a1 2251 TGGGGAAACAGGTCTCCCAGGAATGCCCGGAGCTGATGGACCCCGGGG 2300
proCol5a1 T7 359 TTGGAAAACAGGTCTTCCAGGAATGCCCGGAGCTGATGGACCTCCAGGT 408
* ** ** ** ** * ** ** * ** ** * ** ** * ** ** * ** ** * ** ** *

Rat Col 5a1 2301 CACCCCTGGCAAAGAAGGTCCTCCAGGAGAGAAAGGAGGCCAGGGTCCTCC 2350
proCol5a1 T7 409 CACCCGGGTAAAGAAGGACCATCAGGGGAGAAAGGAACCCAGGGACCTGC 458
***** ** ** ** * ** ** * ** ** * ** ** * ** ** * ** ** *

Rat Col 5a1 2351 CGGCCCCCAGGGTCCCATTGGCTACCCAGGTCCACGAGGAGTCAAGGGGG 2400
proCol5a1 T7 459 AGGGCCACAGGGGCCACTCGGATACCCAGGATCACGTGGTGTAAAGGGCG 508
** ** ** * ** * ** ** * ** * ** ** * ** ** * ** ** *

Rat Col 5a1 2401 CAGATGGCATCCGAGGGCTGAAGGGCACCAAGGGAGAGAAGGGTGAAGAC 2450
proCol5a1 T7 509 CTGATGGAATTCGTGGTTTAAAGGGAAAAAAAAAAAAA 548
* ** ** * ** ** * ** ** * ** * ** * ** * ** * ** *
```

Appendix D3

Nucleotide sequence of proCol5a2 compared to Hs proCol5a2

HsCol5a2	1001	GAGGGCCCTCCTGGTGAGCCTGGCCACCTGGCCCTCCGGGTCCCCCTGG	1050
Col5a2 Sp6	1	GCTGGTCCGGT	11
		* * * *	
HsCol5a2	1051	CCACCTTACAGCTGCTCTTGGGGATATCATGGGGCACTATGATGAAAGCA	1100
Col5a2 Sp6	12	CCGGAATTCCTGGGATATCGTCGAC-CCACGCGTCCGTATGATGATACTA	60
		** * * * * *	
HsCol5a2	1101	TGCCAGATCCACTTCCTGAGTTTACTGAAGATCAGGCGGCTCCTGATGAC	1150
Col5a2 Sp6	61	TGCCAGAACCCTTCCTGAATTCACAGACGATGAAGCTGCT--GGTGAT	107

HsCol5a2	1151	AAAAACAAAACGGACCCAGGGGTTTCATGCTACCTGAAGTCACTCAGTAG	1200
Col5a2 Sp6	108	GGAACAAAACGGACCCAGGTGTTTCATGCTACAATAAAATCTTTAAGCAG	157

HsCol5a2	1201	TCAGATTGAAACCATGCGCAGCCCGATGGCTCGAAAAAGCACCCAGCCC	1250
Col5a2 Sp6	158	TAAGATTGCAACGATGCGTAGTCTGTATGGATCCAAGGAGAACCAGCAC	207
		* * * * *	
HsCol5a2	1251	GCACGTGTGATGACCTAAAGCTTTGCCATTCCGCAAAGCAGAGTGGTGAA	1300
Col5a2 Sp6	208	GCACATGTGATGATCTAAACTTTGTCACTCTAATAAGAAGAGTGGTGAA	257

HsCol5a2	1301	TACTGGATTGATCCTAACCAAGGATCTGTTGAAGATGCAATCAAAGTTTA	1350
Col5a2 Sp6	258	TATTGGATAGATCCTAATCAGGGATGCGCAGAAGATGCTGTAAAGGTTTT	307
		** * * * *	
HsCol5a2	1351	CTGCAACATGGAACAGGAGAAACATGTATTTACGCAAAACCCATCCAGTG	1400
Col5a2 Sp6	308	CTGTAACATGGAAACTGGAGAAACGTGTATATCTGCAAAATCCATCCTCAA	357

HsCol5a2	1401	TACCACGTAAACCTGGTGGGCCAGTAAATCTCCTGACAAT--AAACCT	1447
Col5a2 Sp6	358	TACCTCGAAAAACATGGTGGACTGGTAAAGGAGTGGATAACTTCAAGCCA	407
		**** * * * * *	
HsCol5a2	1448	GTTTGGTATGGTCTTGATATGAACAGAGGGTCTCAGTTCGCTTATGGAGA	1497
Col5a2 Sp6	408	ATATGGTATGGCGCTGATATGAACAAAGGCTCTCAATTTGTCTATGGAGA	457
		* * * * *	
HsCol5a2	1498	CCACCAATCACCTAATACAGCCATTACTCAGATGACTTTTTTGCGCCTTT	1547
Col5a2 Sp6	458	CACTCAATCACCGAATACAGCTGTCACCCAAGTACCTTCCCTCCGCCTTC	507
		* * * * *	
HsCol5a2	1548	TATCAAAAGAAGCCTCCGAGAATCATTACATCTGTAAAAACAGTGTA	1597
Col5a2 Sp6	508	TGTCAAAGGAAGCTTCTCAGAATATAACCTACTATGTAAAAATGCTGTT	557
		* * * * *	
HsCol5a2	1598	GGATACATGGACGATCAAGCTAAGAACCTCAAAAAAGCTGTGGTTCTCAA	1647
Col5a2 Sp6	558	GGATACTTGGATGAAAAACACAAAACCTGAAAAAGCTGTCA	600

Appendix D4

Alignment of proCol11a1 versus corresponding region of Hs proCol11a1

HsCol11a1	4851	AAAGGGTAACAAAGGCTCTACTGGACCCGCTGGCCAGAAAGGTGACAGTG	4900
Col11a1 Sp6	1	GATATCGTCGACCCACGCTCCGCGGACGCTGGGTGGTG	40
		* * * * *	
HsCol11a1	4901	GTCTTCCAGGCGCTCCTGGGCTCCAGGTCCACCTGGTGAAGTCATTGAG	4950
Col11a1 Sp6	41	GTCTTCCAGGCGCTCCTGGTCCACCTGGGCTCCTGGGGAAGTGATCCAA	90

HsCol11a1	4951	CCTTTACCAATCTTGTCTCCAAAAAACGAGAAGACATACTGAAGGCAT	5000
Col11a1 Sp6	91	CCATTGCTATCCAGTCACCAAGGAAGACCCGAAGATCTTCAGATGATAT	140

HsCol11a1	5001	GCAAGCAGATGCAGATGATAATATTCTTGATTACTCGGATGGAATGGAAG	5050
Col11a1 Sp6	141	GATGGCAGATGCTGGAGACTACACCAT---GTACGAAGATGAAAAGAACC	187
		* * * * *	
HsCol11a1	5051	AAATATTGGTTCCCTCAATTCCCTGAAACAAGACATCGAGCATATGAAA	5100
Col11a1 Sp6	188	AGCTATTTAATTCCTGGATAAACTAAAAGAACAAGTCAGAACATGAAA	237
		* * * * *	
HsCol11a1	5101	TTTCCAATGGGTACTCAGACCAATCCAGCCGAACTTGTAAGACCTGCA	5150
Col11a1 Sp6	238	TTTCCCACAGGAGCGTGACACCCAGCAAGGACATGCAAGACCTGCA	287

HsCol11a1	5151	ACTCAGCCATCCTGACTTCCAGATGGTGAATATGGATTGATCCTAACC	5200
Col11a1 Sp6	288	GCTTAGTCACCCAAAATTACCGATGGTGAATACTGGATTGATCCAAACC	337
		* * * * *	
HsCol11a1	5201	AAGGTTGCTCAGGAGATTCTTCAAAGTTTACTGTAATTTACATCTGGT	5250
Col11a1 Sp6	338	AGGCTGCTCCAGTGATTCAATGAAAGTTTACTGCAATTTACAGCTGGA	387
		* * * * *	
HsCol11a1	5251	GGTGAGACTTGCATTTATCCAGACAAAAATCTGAGGGAGTAAGAATTTC	5300
Col11a1 Sp6	388	GGAGAGACTTGTATATATCCAGACAAAGAAATCAGAAGGAGTGAGGATATC	437
		* * * * *	
HsCol11a1	5301	ATCATGGCCAAAGGAGAAACCAGGAAGTTGGTTAGTGAATTTAAGAGGG	5350
Col11a1 Sp6	438	ATCCTGGCCTAAAGAAAACCTGGCTCATGGTTAGTGAGTTCAAGAGAG	487

HsCol11a1	5351	GAAACTGCTTTTCACTTAGATGTTGAAGGAAATTCATCAATATGGTG	5400
Col11a1 Sp6	488	GAAAGCTGCTGTCTTATGTGGACACTCAGGGTAACACCATTAACCAAGTA	537

HsCol11a1	5401	CAATGACATTCCTGAAACTTCTGACTGCCTCTGCTCGGCAAAATTTAC	5450
Col11a1 Sp6	538	CAATGACCTTCC	550

HsCol11a1	5251	GGTGAGACTTGCATTTATCCAGACAAAAATCTGAGGGAGTAAGAATTTC	5300
Col11a1 T7	1	AGGAGTGAGGATATC	15

HsCol11a1	5301	ATCATGGCCAAAGGAGAAACCAGGAAGTTGGTTAGTGAATTTAAGAGGG	5350
Col11a1 T7	16	ATCCTGGCCTAAAGAAAACCTGGCTCATGGTTAGTGAGTTCAAGAGAG	65

HsCol11a1	5351	GAAACTGCTTTTCACTTAGATGTTGAAGGAAATTCATCAATATGGTG	5400
Col11a1 T7	66	GAAAGCTGCTGTCTTATGTGGACACTCAGGGTAACACCATTAACCAAGTA	115

HsCol11a1	5401	CAATGACATTCCTGAAACTTCTGACTGCCTCTGCTCGGCAAAATTTAC	5450
Col11a1 T7	116	CAATGACCTTCCTGAGACTCCTCAGTCAATCTGCCGGCAAAACTTCAC	165

HsCol11a1	5451	CTACCACTGTATCAGTCAGCAGCCTGGTATGATGTGTATCAGGAAGTT	5500
Col11a1 T7	166	CTATACTTGTCAACACTCTGTTGCTTGGCATGACAATGTGGCCAGCACT	215

HsCol11a1	5501	ATGACAAAGCACTTCGCTTCCTGGGATCAAATGATGAGGAGATGTCCTAT	5550
Col11a1 T7	216	ACAACAAGGCCCTAAGGTTCTTGGATCCAATGATGAGGACATTTCTTAT	265
		* * * * *	
HsCol11a1	5551	GACAATAATCCTTTTATCAAAACACTGTATGATGGTTGTACGTCCAGAAA	5600
Col11a1 T7	266	GATAACAACCCATACATCAGACCTTTGCATGATGGCTGTGCGCTGAAGAA	315

HsCol11a1	5601	AGGCTATGAAAAGACTGTTCATTGAAATCAATACACCAAAATTTGATCAAG	5650
Col11a1 T7	316	AGGATATGCACGAACAGTCCTGGAAATTAACACCTGAAGATAGATCAGC	365

HsCol11a1	5651	TACCTATTGTTGATGTATGATCAATGACTTTGGTGATCAGATCAGAAG	5700
Col11a1 T7	366	TGCCATTGCAGATGTAATGATCAGTGACTTTGGGGAGCAGCACCAGAAA	415
		* * * * *	
HsCol11a1	5701	TTGCGATTGAAGTTGGTCTGTTTGTCTTGGCTAAGATTAAAGACAA	5750
Col11a1 T7	416	TTTGGAATTTGAGGTCGGACCGGTCTGCTTCTTAGGTTGAGACTGGAATGT	465

HsCol11a1	5751	AGAACATATCAATCAACAGAAAATATACCTTGTCGCCACCAACCCATTT	5800
Col11a1 T7	466	GGACAAG--CAGCTTAATAAAAAAAAAAAAAA	496
		* * * * *	

Appendix E

Alignment of Full length Xbgn to Xbgn EST nucleotide sequence

Xbgn (AB037269)	1	CTCACATACAAGACCCTGCAGCAGGCAGCCCTGTTCTAGTCAGTACAAGG	50
Xbgn EST	1		0
Xbgn (AB037269)	51	GGGGCGGTTAGGGTGGATGCCTCCTTTTGTACAGAAGAAAAGTCAGAAA	100
Xbgn EST	1		0
Xbgn (AB037269)	101	AGTGAGGAGGAGAAGAAGAACTGGAAGATTTTGTTTAAAAAAGTGAAAA	150
Xbgn EST	1		0
Xbgn (AB037269)	151	AAAAAAACGTTCTGAACCAGAGACTGTGATATTATAAACGGGACGCAAG	200
Xbgn EST	1		0
Xbgn (AB037269)	201	TTTATTGCCTGGAAGGAGCCTTTGTATATTGAAAGAGAAAAGAGAGAGAG	250
Xbgn EST	1		0
Xbgn (AB037269)	251	ATATCAACACCAACAATGAAGGTATTGCTTCTCTTATGCTCCTGTATCTT	300
Xbgn EST	1	START	0
Xbgn (AB037269)	301	AGTCATCCATGCCCTGCCTTTTGAACAGAGAGGATTCTGGGATTTTTC	350
Xbgn EST	1		0
Xbgn (AB037269)	351	TGGACGATGGAATGGCTATGATGAAAGATGAGGAGGCCTCGGGAGTAGGT	400
Xbgn EST	1		0
Xbgn (AB037269)	401	CCTATACCTACTGAGAGTATTCCAGATGTTGGTTTGCCACCTATGGATT	450
Xbgn EST	1		0
Xbgn (AB037269)	451	ATGCCCATTTGGGTGTCAATGTCATTTGCGTGTGGTGCAGTGCTCTGATT	500
Xbgn EST	1		0
Xbgn (AB037269)	501	TGGGTCTCACGTCTATTCCTAAAAACCTTCCGAAAGACACCACTCTCTTG	550
Xbgn EST	1		0
Xbgn (AB037269)	551	GATCTACAGAACAAACAAATCACGGAAATCAAAAGGACGACTTCAAGGG	600
Xbgn EST	1		0
Xbgn (AB037269)	601	ACTCACCAACCTGTATGCTTTGGTTATAGTCAACAACAAGATATCGAAAA	650
Xbgn EST	1		0
Xbgn (AB037269)	651	TCAATGAAAAGGCCTTTGAGCCCTGCAGAAGATGCAGAACTGTACATC	700
Xbgn EST	1		0
Xbgn (AB037269)	701	TCTAAAAACAACCTCGAGGAAATACCCAAAAACCTTCCCAAGTCTCTGGT	750
Xbgn EST	1		0
Xbgn (AB037269)	751	GGAGCTGCGCATTCACGAAAACAAATTAAGAAGGTCCCAAAGGCGTCT	800
Xbgn EST	1		0
Xbgn (AB037269)	801	TCAGTGGACTGAAGAACATGAACTGTATTGAAATGGGAGGAAACCTCTG	850
Xbgn EST	1		0

Xbgn (AB037269)	851	GAAAAATGGAGGCATTGAGGCTGGCGCATTGATGGCTTGAAGCTGAAC	900
Xbgn EST	1		0
Xbgn (AB037269)	901	CTTGCGTGTTCGAGGCAAACTGTCCGGCATCCCAAAGGTCTTCCAA	950
Xbgn EST	1		0
Xbgn (AB037269)	951	GCACCCTAAATGAGCTTCACCTTGACAACAACAGATCCAGGCTATTGAA	1000
Xbgn EST	1		0
Xbgn (AB037269)	1001	AAGGAAGACCTTAGCCAGTATGCTAGCCTGTACAGGCTGGGTTTAGGACA	1050
Xbgn EST	1	TGCTAGCCTGTACAGGCTGGGTTTAGGACA	30

Xbgn (AB037269)	1051	CAACAACATCCGCATGATAGAAAACGGCAGCTTATCTTTTATGCCTGTTC	1100
Xbgn EST	31	CAACAACATCCGCATGATAGAAAACGGCAGCTTATCTTTTATGCCTGTTC	80

Xbgn (AB037269)	1101	TGAGGGAGCTTCACCTGGACAAACAACTCTCTAAAGTGCCCCAGGG	1150
Xbgn EST	81	TGAGGGAGCTTCACCTGGACAAACAACTCTCTAAAGTGCCCCAGGG	130

Xbgn (AB037269)	1151	CTGCCAGACATGAAATTACTGCAGGTTGTGTATCTGCACTCCACAACAT	1200
Xbgn EST	131	CTGCCAGACATGAAATTACTGCAGGTTGTGTATCTGCACTCCACAACAT	180

Xbgn (AB037269)	1201	CACCCAGGTTGGCGTCAATGACTTCTGCCCTATAGGATTCGGAGTGAAC	1250
Xbgn EST	181	CACCCAGGTTGGCGTCAATGACTTCTGCCCTATAGGATTCGGAGTGAAC	230

Xbgn (AB037269)	1251	GCGCCTACTACAACGGAATTAGTCTATTTAACACCCCGTGCCCTACTGG	1300
Xbgn EST	231	GCGCCTACTACAACGGAATTAGTCTATTTAACACCCCGTGCCCTACTGG	280

Xbgn (AB037269)	1301	GAAGTGCAGCCGGCCACATTCGCTGCGTCACCGACCGATTGGCTATCCA	1350
Xbgn EST	281	GAAGTGCAGCCGGCCACATTCGCTGCGTCACCGACCGATTGGCTATCCA	330

Xbgn (AB037269)	1351	GTTTGGCAATTACAGGAAGTAAAGAAGGTGGAACGAGATGCAGGAACCCC	1400
Xbgn EST	331	GTTTGGCAATTACAGGAAGTAAAGAAGGTGGAACGAGATGCAGGAACCCC	380

		STOP	
Xbgn (AB037269)	1401	CTTATTGTAAGAGGCATACGACAGACACGAAGAGAGCAGAAGAATCAAAT	1450
Xbgn EST	381	CTTATTGTAAGAGGCATACGACAGACACGAAGAGAGCAGAAGAATCAAAT	430

Xbgn (AB037269)	1451	TTGCAGAAATTACCAATTATGTTTAAACACAAGGGAGTGTAAACCTAAAA	1500
Xbgn EST	431	TTGCAGAAATTACCAATTATGTTTAAACACAAGGGAGTGTAAACCTAAAA	480

Xbgn (AB037269)	1501	ACAATTATAACCCCCCCCC--AACACCCAACTGTAGGTATCTGCCCTGA	1548
Xbgn EST	481	ACAATTATAACCCCCCCCCAACACCCAACTGTAGGTATCTGCCCTGA	530

Xbgn (AB037269)	1549	ATTGCCGTGTCTTACCGTCTTCTGGCACCCTGTATATAGAACATGGAAGG	1598
Xbgn EST	531	ATTGCCGTGTCTTACCGTCTTCTGGCACCCTGTATATAGAACATGGAAGG	580

Xbgn (AB037269)	1599	ATGTGCAGATAGCACTACTCATATGACCACAAATTGTGTAACCCCCACAC	1648
Xbgn EST	581	ATGTGCAGATAGCACTACTCATATGACCACAAATTGTGTAACCCCCACAC	630

Xbgn (AB037269)	1649	TCGCACAATGATACGCTTCTTTTATAGAGACTATGCAACTTGTAAACAGA	1698
Xbgn EST	631	TCGCACAATGATACGCTTCTTTTATAGAGACTATGCAACTTGTAAACAGA	680

Xbgn (AB037269)	1699	AATAGCAGATTTCTAGATCCCCCCTGTACTGTAGTGCAGAGCATGGACA	1748
Xbgn EST	681	AATAGCAGATTTCTAGATCCCCCCTGTACTGTAGTGCAGAGCATGGACA	730

Xbgn (AB037269)	1749	TGCTGGCCATGGGGATTAGAAGACTTTAGCTCATTGTGTTGCTTGAAG	1798

Xbgn EST	731	TGCTGGCCATGGGGATTAGAAGACTTTAGCTCATTGTGTGTGCTTGAAG	780

Xbgn (AB037269)	1799	CTT-----TTGGGTCTTATACCCCTTCCACTCCTTTTACATGCCATCC	1842
Xbgn EST	781	CTTAAGCTTTTGGGTCTTATACCCCTTCCACTCCTTTTACATGCCATCC	830
		*** *****	
Xbgn (AB037269)	1843	CTCACGTAGTCAGTACATAGGCTTCTGCAGTCACACTAACAGCTCACAAG	1892
Xbgn EST	831	CTCACGTAGTCAGTACATAGGCTTCTGCAGTCACACTAACAGCTCACAAG	880

Xbgn (AB037269)	1893	CACCCAGTCTAGCTCTCTGCTGGAGCTACCATATGTCTTCTTTTTCACG	1942
Xbgn EST	881	CACCCAGTCTAGCTCTCTGCTGGAGCTACCATATGTCTTCTTTTTCACG	930

Xbgn (AB037269)	1943	CTCAGATTGCGGAGATTGGTCGCCCGGGTCAAACCTCCTCTTCTTCTC	1992
Xbgn EST	931	CTCAGATTGCGGAGATTGGTCGCCCGGGTCAAACCTCCTCTTCTTCTC	980

Xbgn (AB037269)	1993	CCCAAAGTGCCTCCCGCCATCTAGAGTGTAGATTGCCGGCGGGATGGCAC	2042
Xbgn EST	981	CCCGGACTGCCTCCCGCCATCTGGAATGTAGGTTGCCGGCGGGATGGCAC	1030
		*** *****	
Xbgn (AB037269)	2043	TCAGAGTTTTCGTTTTCCCGAAATCGGCTCAAGTTTCCTTGTGAGGCAA	2092
Xbgn EST	1031	TCAGAGTTTTCGTTTTCCCGAAATCGGCTCAAGTTTCCTTGTGAGGCAA	1080

Xbgn (AB037269)	2093	CTTCGGAAAAAGACATGCCATCCCGCCAGCGATTATGGTCTGGCCGACG	2142
Xbgn EST	1081	CTTCGGAAAAAGACATGCCATCCCGCCAGCGATTATGGTCTGGCCGACG	1130

Xbgn (AB037269)	2143	GGAGGTAGTTTGGGGAGATTAGTCGCCCCGAAGAAGAGGAGGTTTGTCCC	2192
Xbgn EST	1131	GGAGGTAGTTTGGGGAGATTAGTCGCCCCGAAGAAGAGGAGGTTTGTCCC	1180

Xbgn (AB037269)	2193	CTGGGCGACTGGTCTCCCCGAATCTGAGCGTGTGCCCTGACCTGAATAT	2242
Xbgn EST	1181	CTGGGTGACTGGTCTCCCCGAATCTGAGCGTGTGCCCTGACCTGAATAT	1230

Xbgn (AB037269)	2243	CATTCCAATTCTGTAGTCTCCATAATCCCTTTTAACCCCTTCAGTACTGT	2292
Xbgn EST	1231	CATTCCAATTCTGTAGTCTCCATAATCCCTTTTAACCCCTTCAGTACTGT	1280

Xbgn (AB037269)	2293	CTGAGACACATTGCAGAGCAATGGCTTGTGACCCTCACATCACCAGATA	2342
Xbgn EST	1281	CTGAGACACATTGCAGAGCAATGGCTTGTGACCCTCACATCACCAGATA	1330

Xbgn (AB037269)	2343	GGTCTTGCCAGCATATCGGTTTCATCATATTTAACCCCTTTATTTAAGTT	2392
Xbgn EST	1331	GGTCTTGCCAGCATATCGGTTTCATCATATTTAACCCCTTTATTTAAGTT	1380

Xbgn (AB037269)	2393	ATGATGACCACCTTTAGCTTTAGGAACGATTAAGCAGCCTGGAATAATGGC	2442
Xbgn EST	1381	ATGATGACCACCTTTAGCTTTAGGAACGATTAAGCAGCCTGGAATAATGGC	1430

Xbgn (AB037269)	2443	ATAGCACTTTTATTTTGTTCACAAATTCATGCAATTTTATTTTAAAG	2492
Xbgn EST	1431	ATAGCACTTTTATTTTGTTCACAAATTCATGCAATTTTATTTTAAAG	1480

Xbgn (AB037269)	2493	TAGAAAATCCAATTTCTCCTGCCTCAAGCATCGCTGGGAATATGCTTGG	2542
Xbgn EST	1481	TAGAAAATCCAATTTCTCCTGCCTCAAGCATCGCTGGGAATATGCTTGG	1530

Xbgn (AB037269)	2543	TCGCATTGCCACCTACAGGAGAAGAAAGGGCACTGCGTGGTTTATAGGTG	2592
Xbgn EST	1531	TCGCATTGCCACCTACAGGAGAAGAAAGGGCACTGCGTGGTTTATAGGTG	1580

Xbgn (AB037269)	2593	TCCAAAAATGTTTCTTTTAAATATAAATTCATTAAATAGTCTAGAAAGC	2642
Xbgn EST	1581	TCCAAAAATGTTTCTTTTAAATATAAATTCATTAAATAGTCTAGAAAGC	1630

Xbgn (AB037269)	2643	AGAGGGTGATTCTACTGTATGTGGTACAGTGGAAACCACACACAACAG	2692
Xbgn EST	1631	AGAGGGTGATTCTACTGTATGTGGTACAGTGGAAACCACACACAACAG	1680

```

*****
Xbgn (AB037269) 2693 CAAGTCCTTTGGCATTCCATTTGAAAGAGCAAAGATCAGAGACGGCAAGC 2742
Xbgn EST 1681 CAAGTCCTTTGGCATTCCATTTGAAAGAGCAAAGATCAGAGACGGCAAGC 1730
*****

Xbgn (AB037269) 2743 TTAAAAGAAGACCGCAGCATTTCTGTAGATGCCAAAGAAACCCTCCCATT 2792
Xbgn EST 1731 TTAAAAGAAGACCGCAGCATTTCTGTAGATGCCAAAGAAACCCTCCCATT 1780
*****

Xbgn (AB037269) 2793 TATGCAGCTTTTGGGTATAAATGGAAAAGTGTATTTTAGGAAAGGTAGC 2842
Xbgn EST 1781 TATGCAGCTTTTGGGTATAAATGGAAAAGTGTATTTTAGGAAAGGTAGC 1830
*****

Xbgn (AB037269) 2843 AGTAATTTTCAGTTATTGTATATCCAGTGTGATGTGAGGAAGGAATACTTA 2892
Xbgn EST 1831 AGTAATTTTCAGTTATTGTATATCCAGTGTGATGTGAGGAAGGAATACTTA 1880
*****

Xbgn (AB037269) 2893 CAAACTGTGCTGAGCATTTTCTTTTCTTTAGCATCAAATAATGAAGCATTA 2942
Xbgn EST 1881 CAAACTGTGCTGAGCATTTTCTTTTCTTTAGCATCAAATAATGAAGCATTA 1930
*****

Xbgn (AB037269) 2943 AAATATTGTCCTTTTCTC 2960
Xbgn EST 1931 AAATATTGTCCTTTTCTCCCCAAAAAAAAAAAAA 1965
*****

```

Nucleotide and Amino acid sequence of Laminin α -like

CGGGAGGACTCTTCTGTGTTTCATCCCAAAGCAGTGAAGGGAGCTCTACGATTGGTGATCTTCCAGCAGCC
R E D S S C V H P K A V K G A L R F G G S S S >
100 150

ACCTAGAATTCACCAGACTCCCACCTATCTTTCAGAGACAGGTTTCACTTTAGCTTAGAGATCCGCTCGAACACCT
H L E F T R L P P I F R D R F H F S L E I R L N T>
200

CCAATGGGTTGATCTTTTTACGCCCGGGATGAGAGAAGCAGCTCTCCCTTTTTCTGTATGTGTCCAATGGTCGCA
S N G L I F Y G R D E R S S S S S L F L Y V S N G R>
250 300

TGTGATTGCATCTGGTTATCATAGGGAAGTCACTGCGTTTGAGGAGCAAAAGAAAAGTACAACGATGGACTCTGGC
I V L H L V I I G K S L R L R S K E K Y N D G L W>
350

ACACGGTATTCTTTGGCAAAGAGAAGAATAAGCTGCACTTGGTCATAGATGGCATTAAAGCCCAGAGCAGCATGG
H T V F F G K E K N K L H L V I D G I K A Q S S M>
400 450

TGTATCCTGGAGTTAATCATCTTTAACTGGTCCCGTCTTCATCGGAGGTCTCCCACCGCTCATCAGAAGACCTG
V Y P G V K S S L T G P V F I G G L P P L I R R P>
500

ACATTCCTGATGTGCTCACTCAGCAGTTTTCACGGATGTCTCAGGAACITGAAGTTGGATGGGAAAGCACTTAACC
D I P D V S L S S F H G C L R N L K L D G K A L N>
550 600

CCCCAAATAAAATCATGGGGGTCAACCCAGTGCTATGAAGGGCTCACCGAACAAGGATTTTTCTTCTCTGAGGGTG
P P N K I M G V T Q C Y E G L T E Q G F F F S E G>
650

GATTCTTAAATTAGACCATGAAGTGGATATTGGGCAGAGCTTGAAGTGAAACTGGAGATACGTCCAATCCATC
G F L K L D H E V D I G Q S L E V K L E I R P I H>
700 750

ATTACGGCCTCCTTTTCCATGTGGGAAGTGAAGGGGCCACTCTGTCAACCTTTCCATGACTGAGGGGAAGGTCA
H S G L L F H V G T E E G H S V T L S M T E G K V>
800

GTGTGTCTGTAAATGCAGGAGCAAGTGAATATTCACATCTGTGAAATTCACCAACCCCTCTGTGATGGACAGT
S V S V N A G A S E Y S T S V K F P Q P L C D G Q>
850 900

GGCACACAGTTGCAGTGACCAAGGTTAGTAATGTTATCCAGCTGGATGTGGATACAGAGGGAACCATGCTGTAG
W H T V A V T K V S N V I Q L D V D T E G N H A V>
950

GAGCAAGCCAGACAGAGCCAGCAGGAGCAGGGGGACTCTCTACATCGGAGGAGTGCCTGGAACATCTACAGCC
G A S Q T E P S R S R G T L Y I G G V P G N I Y S>
1000 1050

CGAGGAAACCGCACTCATCTTATCAGGCTGCATGAGGAATCTTGTATTAAACCGAAAGCAGTTGACGCTCCCA
P R K P H S S Y R G C M R N L V I N R K A V D V S>
1100

AACCTGGGACTTTCTGTAGGATCAGCCGGACCAACGTCTGCCCGTCCCTGTGATCCCAATGCACAAAAATGGAAGG
K P G T F V G S A G T N V C P S L *
1150 1200

ACTGTACAAACAAGCGAGCCAGCCTACTGCTGAAACCTTTGAGACAGATGGTGTACACTGTGCAAGTACTGCGT
1250

GTGGCTTAGGGGTAATGTGACAAGAAGGCCCTTAAAGCCCACTTGGAAATATTGGGTGCTTGGAAAATAGCATATG
1300 1350

GTTACTGGATTAAACTCTTTTACACAGCAGAACCCTAACAGCGAGCTAACCCACCCCTTATTTTAAAGTTCCAA
1400

ACAGCATTATTCCACTCTGATACTATCATGCTTCTAATCAACATCCAGGGAAGGGGGTGGGTACCTTTAAAAA
1450 1500

CTGTATCTGGGGAAGAAGCCCAAAGCAGTGTTAGTCAGTTTTGACAGTTTGATTATTCCAAGTCTGTCCCCAGT
1550

GGGTACACAAATGTTACACAATACGAAGGAATCTACGCAACTCTCTATTGGACCAAAAAGCTATAGACTGTTGA
1600 1650

AAACATTACGCTTACAGTTGACTTCTGGTTTAAAGGGGATGTTCAACTTTGAGATAACTTATTATGATGCAGAG
1700

GGTGACAATTGCAATTGGTTTCCATTATTTATGGTTATTTAGCTTTTATTTCAGCAGCTTTCCAGTTTTCAGCAG
1750 1800

TCTGGTTACTAGGGTCCAAATAACTAGCAACCATGCAGGAATTTATATGTAGACTGGAATATGAACAGGAGAGGAT
1850

CTAAATAGAAAGAGGAGTAATAAAAAGTAGCAATAACAATAAATGTGTAGCCTTAATGAGCATCTGTTTTAAATT
1900

GGGGTCAGTGGCTCCCAATTGGAAGCTGGAAGAGTGAGAAGAAGGCAAAATAATAGAAAAAAAAAAAAAAAAAAAA

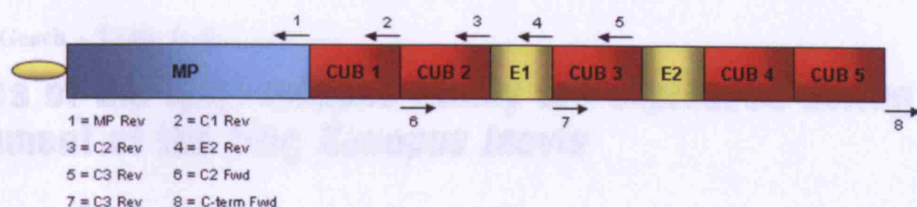
Appendix G

Dorso-Anterior Index (DAI)

Embryos were scored for their level of ventralisation or dorsalisation based on the DAI scale of Kao and Elinson (1988). Normal embryos are scored as 5, ventralised embryos are scored between 4 and 1 (1 being the most ventralised phenotype) and dorsalised embryos scored between 6 and 10 (10 being the most dorsalised). Kao and Elinson described the DAI at much later stages than those analysed in this study, below is a description of the criteria used in this study for ventralised embryos at St. 26/27 based on their original description.

- 5** Normal embryo
- 4** Reduction in head size and enlargement of ventral-posterior, however all structures still appear to be present
- 3** Further reduction in head size, but coupled with smaller cement gland and no obvious development of the eye, further enlargement of the ventral-posterior.
- 2** No obvious head or cement gland, extremely large ventral-posterior region, still able to establish anterior from posterior
- 1** No head and now difficult to determine anterior from posterior, can still make out dorsal-ventral axis.

Appendix H



Deletion construct	Fwd Primer	Rev Primer
XldC1	C-term Fwd	C1 Rev
XldC2	C2 Fwd	MP Rev
XldC1-2	C-term Fwd	C2 Rev
XldC1-e	C-term Fwd	E1 Rev
XldC1-3	C-term Fwd	C3 Rev
XldC3-5	C3 Fwd	MP Rev
XldC1+3-5	C3 Fwd	C1 Rev

Primer	Sequence
C-term Fwd	ttcaggatgccttcac
C2 Fwd	tgcgaggagacatcaaaaaagactctgg
C3 Fwd	tgtggaggttttattacacaac
C1 Rev	tatagcttcatatgtctcaa
C2 Rev	tttaaagaaattggcagaaa
E1 Rev	tgcgcttcacaactttt
C3 Rev	ttcatccttatcagagaaaaag
MP Rev	tgctgggcatttatagatttc

